DISCOVERY AND ANALYSIS OF COMPONENTS OF NONSENSE MEDIATED DECAY IN *DROSOPHILA*

by

Kimberly Ann Frizzell

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Human Genetics

The University of Utah

May 2013

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Kimberly Ann Frizzell**

has been approved by the following supervisory committee members:

Mark Metzstein	, Chair	**12/13/2012** Date Approved
Carl Thummel	, Member	**12/13/2012** Date Approved
Brenda Bass	, Member	**12/13/2012** Date Approved
Charles Murtaugh	, Member	**12/13/2012** Date Approved
Gillian Stanfield	, Member	**12/13/2012** Date Approved

and by **Lynn Jorde**, Chair of the Department of **Human Genetics**

and by Donna M. White, Interim Dean of The Graduate School.

ABSTRACT

Nonsense mediated decay (NMD) is a cellular surveillance mechanism originally identified for its role in targeting mRNAs containing nonsense codons for rapid degradation. Nonsense codons, or premature termination codons (PTCs), are a common result of genomic mutation, alternative splicing, and programmed frameshifts that bring a premature termination codon into the reading frame. Degradation of the mRNA encoding PTCs is thought to prevent the translation and build-up of potentially harmful truncated proteins within the cell. The phenomenon of NMD is evolutionarily conserved, as destabilization of PTC-containing mRNA occurs in all eukaryotes examined. *Trans*-acting factors responsible for the destabilization of PTC-containing mRNA were identified in a series of informational suppressor screens performed in yeast and *C. elegans*. Combined, these screens identified seven genes that are conserved throughout eukaryotes. Homologs of genes identified in the original screens have been analyzed biochemically to determine how these *trans*-acting factors can recognize a premature termination codon. Two models of PTC-recognition have resulted from these biochemical studies; the primary model suggests a PTC causes aberrant translation termination that mimics an abnormally long 3'UTR. The second models suggests factors remain associated with the mRNA after intron splicing, and these factors recruit the NMD machinery to the mRNA, marking it for degradation. However, while analyzed biochemically, many questions remain about the mechanism of NMD. Have all the genes

required for NMD been identified? What are the roles of NMD pathway genes *in vivo*? How does NMD differentiate a native termination codon from a premature termination codon? Work presented in this thesis focuses on these questions. Here we present a forward genetic screen in *Drosophila*, which isolated alleles of known and potentially novel components of the NMD pathway. We use alleles isolated from this screen to investigate the mechanism of NMD function *in vivo*. Finally, we analyze *cis*-acting sequences involved in NMD targeting in *Drosophila*.

This thesis is dedicated to my husband, Jesse.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Mark Metzstein. I am lucky to have been trained by such a thoughtful and intelligent person. Mark is a great scientist who truly cares about his work, but cares equally about each of his students and makes sure we fully understand the science. I appreciate the hours of help that Mark has provided improving my presentations, writing, and thought process. I will be a better scientist because of the time I spent with Mark. Finally, Mark has been a good friend, someone you wanted to talk both science and life with. I will always be grateful for my time spent in his lab.

I would like to thank my committee, Carl Thummel, Brenda Bass, Gillian Stanfield, and Charlie Murtaugh. Their support and guidance throughout the years has been instrumental in my development as a scientist. Committee meetings were a source of focus, encouraging me to consolidate my efforts on multiple projects into a publishable manuscript. The personal feedback they provided me when thinking about future career options was invaluable.

I would like to thank the Thummel, Yandell, and Stanfield labs. We spent the first few years as a small lab and were granted the opportunity to participate in joint lab meetings with the Thummel lab. These early experiences were essential in my development as a scientist as they provided me with both an example of good science and feedback on my own work. Even though we outgrew those lab meetings, we still

continue to be good friends with both new and old members of the Thummel lab. The Yandell lab has been a wonderful source of bioinformatic and statistical support, as well as wonderful friends and neighbors. For years we have also spent days holed up in the microscope room with members of the Stanfield lab. You can't spend hours in a small dark room with someone without becoming their friend. Thanks Stanfield lab.

I would like to thank the Metzstein lab, you guys have been way more than lab-mates. All the people in this lab have been wonderful to work with; the atmosphere of this lab made me want to come to work. I'm so lucky to have had the opportunity to spend the years with you, Tiffani Jones, Alex Chapin, and Jon Nelson. I think of you guys as family, thanks for putting up with me with so much patience and understanding.

I would like to thank my in-laws, Brian, Julie, Jordan and Kari, and Ben for all their support during this long process. I am so happy to be part of this family. I would also like to thank my sister Kirsten, who has been supportive of me during this process and has provided me with a wonderful distraction from work. I would also like to thank Deborah Cutler who is one of the most thoughtful and caring people I've ever met.

To all the friends and colleagues who have helped along the way but would require several pages to list, thank you.

Finally I would like to thank my husband, Jesse. Jesse has provided me with constant love and support, he is my best friend. Without him I would never have gotten to where I am today. Also, he has provided me with the two most wonderful dogs in the whole world, Kaikoura and Lyra, my stress relievers.

CHAPTER 1

INTRODUCTION

Regulation of RNA metabolism is a key mechanism in controlling proper levels of gene expression. In particular, modulation of RNA degradation rates provides an integral regulatory step in maintaining steady state message levels. The more rapidly a transcript is degraded the less protein that can be made from that transcript. Thus, RNA degradation provides an additional layer of regulation of the quantity of protein that can be translated at a given time or location (Besse and Ephrussi, 2008). Several RNA degradation mechanisms function in targeting specific transcripts for rapid decay. This specificity is necessary to avoid nonspecific alterations in gene expression.

Multiple pathways that identify transcripts for degradation have been investigated. Some of these pathways, including the miRNA and siRNA pathways, use a template-based homing mechanism to down-regulate specific transcripts based on sequence (Okamura, 2012). Others, including nonstop (Vasudevan et al., 2002) and no-go decay (Harigaya and Parker, 2010), have a more general surveillance role, seeking out general defects without discretion as to the particular transcript being targeted. The nonsense mediated mRNA decay (NMD) pathway is another such surveillance mechanism (Kervestin and Jacobson, 2012). NMD targets a wide range of mRNAs for rapid degradation and is unique in targeting transcripts containing premature termination

codons (PTCs). Premature termination codons can result from a variety of mechanisms including genomic mutations, frameshifts, and alternative splicing which bring a PTC into the open reading frame (McGlincy and Smith, 2008). An important feature of NMD is that it is able to differentiate PTCs from the natural termination codons present at the end of all coding sequences; a difficult task as PTC location is inherently variable since it may occur anywhere in the transcript. NMD is thought to prevent the translation of potentially harmful truncated proteins, which would result from the presence of a PTC. Current research in NMD focuses on several questions; 1) Have all the genes required for NMD been identified? 2) How does NMD differentiate a natural termination codon from a premature termination codon? 3) What roles does the NMD pathway play *in vivo*? This thesis describes my work investigating the NMD pathway using *Drosophila melanogaster*.

The phenomenon of NMD was originally observed in yeast (Losson and Lacroute, 1979), was later shown to affect RNA metabolism in humans (Baserga and Benz, 1988), and is evolutionarily conserved throughout eukaryotes (Culbertson and Leeds, 2003; Gatfield et al., 2003). Two main lines of evidence suggested this conserved mechanism was regulated by *trans*-acting factors. First, PTC-containing transcripts displayed a decreased half-life compared to transcripts containing a natural termination codon, resulting in low steady-state levels of nonsense transcripts (Baserga and Benz, 1988; Belgrader et al., 1994; Chia et al., 1987; Losson and Lacroute, 1979). Second, destabilization of PTC-containing mRNA was disrupted by the translation inhibitor, cycloheximide, (Leeds et al., 1992; Peltz et al., 1993). Additionally, PTC-containing transcripts associate with polysomes (Leeds et al., 1991), suggesting the PTC-dependent

mRNA destabilization was translation-dependent. Together, the low steady state levels, the increased decay rate, translation dependence, and functional conservation of the phenomenon suggested the existence of a conserved *trans*-acting surveillance machinery. However, the gene or genes responsible for the selective destabilization of transcripts encoding nonsense mutations was unknown.

Yeast screens

In 1980, a large-scale analysis of frameshift suppressors was undertaken in yeast (Culbertson et al., 1980; Cummins et al., 1980). Lines that contained a tRNA mutation, *SUF1*, suppressed a histidine mutation, *his4-38,* by decoding a four basepair codon. This allowed for readthrough of the +1 frameshift mutation encoded by the *his4-38* mutation. The +1 frameshift mutation brought a nonsense codon into the reading frame (Donahue et al., 1981); thus making it an NMD substrate. In the presence of a mutation in the NMD pathway, the nonsense codon containing transcript was stabilized, resulting in further suppression of the *his* mutant phenotype. The stabilization of *his-38* transcript was later shown to occur independently of the *SUF1* mutation, suggesting these mutations that "*up*" the suppression of the *his* mutant phenotype play an independent role in mRNA destabilization (Leeds et al., 1991). This analysis identified two separate mutations called *UPF* for "*UP-Frameshift suppressor*" (Culbertson et al., 1980). Three genes named UPF1, UPF2 and UPF3 were eventually identified using these and similar methods (Cui et al., 1995; He and Jacobson, 1995; Lee and Culbertson, 1995; Leeds et al., 1991; 1992). Cloning and further characterization of the UPF genes in yeast revealed all three genes were required to destabilize a broad variety of PTC-containing transcripts,

were dependent on active translation, and were necessary for viability in yeast (Cui et al., 1995; He and Jacobson, 1995; Leeds et al., 1991; 1992). These data suggested all three genes were involved in the same translation-dependent pathway, which functioned to clear the organism of transcripts that could result in truncated proteins.

C. elegans screens

Informational suppressor screens in *C. elegans* also identified mutations in genes affecting stability of PTC-containing transcripts (Cali et al., 1999; Hodgkin et al., 1989). Two mutants were utilized in these screens: *unc-54(r293)* contains a deletion that removes most of the 3'UTR, including the poly-adenylation signal of myosin heavy chain. *lin-29(n546)* introduces a nonsense codon into the coding region of a heterochronic gene. Lines carrying either *unc-54(r293)* or *lin-29(n546)* were mutagenized and recessive suppressor mutations were identified in the F2 generation by reversion (suppression) of the original mutant phenotype. These screens resulted in the isolation of six mutants, each of which suppressed mutations in *unc-54*, *lin-29*, and an additional mutant, *tra-3* (Hodgkin et al., 1989). The mutant suppression was allele-specific, certain classes of mutation were suppressible; however, not all alleles in the same gene were suppressed by this new class of informational suppressors (Hodgkin et al., 1989). On their own, these suppressor mutations exhibited a gross morphological defect, either a swollen bursa in males or protruding vulva in hermaphrodites. Hence they were named *Smg* for "*Suppressor with Morphogenetic effect on Genitalia*" (Hodgkin et al., 1989). The genes identified in this screen included *smg-1*, *smg-2*, *smg-3*, *smg-4*, *smg-5*, and *smg-6* (Hodgkin et al., 1989). The small number of alleles isolated of *smg-4* and *smg-6*, one and

two alleles, respectively, led to the conclusion that this screen was not saturated for genes affecting this pathway. A second screen was performed using a method that prohibited identification of alleles of *smg-1*, *smg-2* and *smg-5*, these provided 90% of the alleles in the original screen. The second screen also utilized a broader variety of chemical mutagens. Using these methods, three alleles of an additional gene, *smg-7*, were identified (Cali et al., 1999). A noncomplementation screen was also performed to identify additional alleles of *smg-6* (Cali and Anderson, 1998). Cloning of the *smg* genes (Cali et al., 1999; Chiu et al., 2003; Cui et al., 1995; Grimson et al., 2004; He and Jacobson, 1995; Page et al., 1999; Serin et al., 2001) showed *smg-2, smg-3,* and *smg-4* were orthologs of the yeast genes UPF1, UPF2, and UPF3, respectively.

C. elegans RNAi reporter screen

In both yeast and *C. elegans*, the UPF and *smg* genes were not required for viability. This raised the question: are there more NMD genes that would have been missed due to lethality upon loss of gene function? To determine whether additional NMD genes were missed due to organismal lethality, an alternative screen was performed (Longman et al., 2007). This screen utilized RNAi to knock down expression of ~86% of the genome of *C. elegans*. Genes required for NMD were identified using an NMD-sensitive fluorescent reporter consisting of a bicistronic transcript that contained GFP fused to a LacZ open reading frame containing a PTC. Under wild-type conditions the reporter transcript was degraded by NMD, resulting in little or no GFP expression. Upon RNAi knockdown of an NMD component, such as *smg-2*, the reporter was no longer targeted for degradation, resulting in GFP expression. Animals carrying the reporter

were fed RNAi in liquid culture and screened for expression of GFP. Two genes weakly stabilized the GFP reporter and also resulted in lethality; these were called *smgl-1* and *smgl-2* for "*Smg-Lethal*" (Longman et al., 2007).

Orthologs of *smgl-1* and *smgl-2* are conserved in eukaryotes: *smgl-1* is homologous to human *NAG/NBAS* and *smgl-2* is homologous to human *DHX34*. Knockdown of *NAG/NBAS* or *DHX34* in HeLa cells using RNAi resulted in stabilization of an NMD reporter. This suggests these two new genes have a conserved NMD function in human cells (Longman et al., 2007). Additionally, homologs of *NAG/NBAS* and *DHX34* are conserved in zebrafish and morpholino knockdown of either gene results in stabilization of a PTC-containing reporter (Anastasaki et al., 2011). However, while these data are consistent with NMD function, the role of these proteins has not been explored mechanistically with respect to other NMD genes, nor have mutants been isolated or tested. Indeed, *DHX34* belongs to a group of RNA helicases that may play a more general role in transcription (Fuller-Pace, 2006). Finally, while a homolog of DHX34 is found in *Drosophila*, no homolog of *NAG/NBAS* is conserved in this organism (Mühlemann et al., 2008).

While the screen described above identified two potential new factors involved in NMD, there are several caveats to the approach. First, RNAi, especially when performed on a large scale, is susceptible to both off-target effects as well as inefficient knock-down (Qiu et al., 2005; Qu et al., 2011); therefore, many genes out of the 86% of the genome screened were likely missed. Second, this screen utilized a bacterial delivery system that results in differential RNAi knockdown in *C. elegans* tissues, so particular genes, such as those required for NMD in the nervous system, would have been missed (Timmons et al.,

2001). Additionally, RNAi knockdown frequently causes lethality; thus, the authors screened both living and dead animals, a method that may have missed genes due to lethality prior to efficient expression of the reporter transgene. Finally, the RNAi approach does not provide the functional data that can be provided by an allelic series, making further analysis of the role of identified genes more difficult.

Protein interaction screens

In addition to *in vivo* screens, two protein interaction screens have successfully identified proteins that could play a role in NMD. UPF2 was identified in a yeast suppressor screen (Cui et al., 1995) and was simultaneously isolated in a yeast two-hybrid using UPF1 as bait (He and Jacobson, 1995). SMG-8 and SMG-9 were identified by coimmunoprecipitation with SMG-1 in HeLa cells (Yamashita et al., 2009). Both proteins were shown to play a minor role in NMD function, likely through regulation of the kinase activities of SMG-1 (Yamashita et al., 2009). However, mutant analysis of the *C. elegans* homolog of *smg-8* demonstrated this gene does not contribute to NMD in *C. elegans* (Rosains and Mango, 2012). Thus, more mutational analysis is necessary to determine whether SMG-8 and SMG-9 contribute to NMD in other organisms.

Forward genetic screen for NMD genes in *Drosophila*

Work to investigate NMD in *Drosophila* has primarily resulted from serendipitous discoveries leading to the isolation of homologs of NMD genes. First, two mutations in *Smg1* were isolated in a forward genetic screen for resistance to the insecticide, cyromazine, in *Drosophila* (Chen et al., 2005). The two alleles of *Smg1* were

shown to have no effect on the levels of PTC-containing transcripts *in vivo*, which led to the conclusion that loss of Smg1 has no effect on NMD *in vivo* (Chen et al., 2005). Another unexpected discovery was made during a forward genetic mosaic screen for genes on the X-chromosome affecting tracheal morphogenesis in *Drosophila* (M. Metzstein, in prep). The screen utilized a fluorescent reporter to positively mark homozygous mutant cells (Lee and Luo, 2001). A class of mutants was recognized that had no morphogenesis defects, but exhibited a drastic increase in fluorescent reporter expression compared to wild-type cells. Mapping and sequencing revealed the defect resulted from mutations in components of the NMD pathway (Metzstein and Krasnow, 2006). Three NMD components are found on the X-chromosome: *Upf1*, *Upf2,* and *Smg1*. Alleles of each of these genes were isolated and analyzed for their role in NMD in *Drosophila*. A single allele of *Smg1* was examined that, like previously analyzed alleles, was viable, but unlike the previous study (Chen et al., 2005) showed a minor transcript-specific NMD defect. This suggests *Smg1* does play a minor role in NMD *in vivo* (Metzstein and Krasnow, 2006). Analysis of alleles of *Upf1* and *Upf2* provided the first *in vivo* data on these genes in *Drosophila*. Two lethal alleles of *Upf1* were isolated, however, the lethality prohibited further analysis in this study. Three alleles of *Upf2* were isolated, two of which were loss of function alleles resulting in lethality. One allele, $Upf2^{25G}$, was a strong hypomorph that resulted in approximately a 10% rate of viability (Metzstein and Krasnow, 2006). Analysis of $Upf2^{25G}$ and the alleles of *Smg1* provided the basis for investigation of NMD in *Drosophila*. While these alleles provided tools to begin analysis of NMD genes in *Drosophila*, there were still no alleles of the *Drosophila* orthologs of the NMD genes *Smg5* and *Smg6* (Gatfield et al., 2003).

Drosophila provides a powerful system to circumvent shortcomings that may have prohibited the original screens from identifying all NMD genes. First, *Drosophila* is a complex organism. Yeast contain orthologs of only three (Cui et al., 1995; He and Jacobson, 1995; Leeds et al., 1991; 1992) of the seven NMD genes identified in *C. elegans* (Cali et al., 1999; Hodgkin et al., 1989), suggesting more factors may be involved in the process as the complexity of the organism increases. Second, the small number of alleles of some of the NMD genes recovered in the *C. elegans* screens (Cali and Anderson, 1998; Cali et al., 1999; Hodgkin et al., 1989), and subsequent identification of genes using RNAi (Longman et al., 2007) and immunoprecipitation (Yamashita et al., 2009), suggests saturation for alleles of NMD genes has not been achieved. Additionally, the knockdown and biochemical approaches did not provide mutants for the identified components, preventing full loss of function analysis of these potentially novel NMD genes. Third, mosaic generation in *Drosophila* provides a way to avoid lethality and developmental defects associated with loss of NMD (Xu and Rubin, 1993). In the original screens mutations were identified by virtue of suppression of a mutant phenotype requiring viability, so lethal mutants would have been missed.

Taking advantage of the mosaic techniques available, we undertook the first forward genetic screen specifically for genes involved in NMD in *Drosophila*. Our screen utilized an NMD-sensitive fluorescent reporter that allowed for efficient identification of NMD genes *in vivo* (Metzstein and Krasnow, 2006). In this system, reporter expression is usually low due to targeting of the reporter transcript by NMD. However, expression increases dramatically in the presence of an NMD mutation due to stabilization of the transcript, resulting in a concordant increase in fluorescent protein

(Frizzell et al., 2012; Metzstein and Krasnow, 2006). We combined the NMD reporter with tissue directed FLP-mediated mitotic recombination to create homozygous mutant cells in an otherwise heterozygous mutant background (Xu and Rubin, 1993). This approach also allowed us to screen F1 animals for mosaic increase in reporter expression. Unlike the original screens in yeast and *C. elegans*, which required homozygous mutants to detect a phenotype, this mosaic approach allows us to bypass any developmental defects associated with NMD genes. Additionally, our approach allows for straightforward detection of NMD mutations. A concern in the original informational suppressor screens was the detection of false positives due to reversions of the mutant phenotypes; however, we avoided this by using a transgenic reporter system instead of a suppressible mutant allele.

Chapter 2 of this thesis describes our forward genetic screen and the resulting alleles of NMD genes. On the X-chromosome all the recovered genes were homologs of mutants previously identified in the *C. elegans* (Cali and Anderson, 1998; Cali et al., 1999; Hodgkin et al., 1989) and yeast screens (Cui et al., 1995; He and Jacobson, 1995; Leeds et al., 1991; 1992), however, potentially novel NMD genes were isolated on the right arm of the third chromosome (3R). Our screen covered ~40% of the *Drosophila* genome, likely saturated the X chromosome, and extensively covered but did not saturate the right arm of the third chromosome (Frizzell et al., 2012). Chapter 3 describes preliminary analysis of alleles isolated on 3R that represent a potentially novel class of NMD mutants.

Mechanism of nonsense mediated decay

Cell culture experiments in mammalian (Nicholson et al., 2009) and *Drosophila* cells (Gatfield et al., 2003), crystallography (Clerici et al., 2009; Kadlec et al., 2004), and mutant analysis in yeast (He and Jacobson, 1995; He et al., 1997) and *C. elegans* (Page et al., 1999) have established protein interactions that occur when a transcript contains a premature termination codon. Studies of interactions between NMD proteins utilized homologs of the seven NMD genes found in the original yeast and *C. elegans* screens. These interactions are generally functionally conserved between orthologs in different species. This results in a protein interaction model of NMD after PTC recognition.

Recognition of a premature termination codon causes Upf1, an ATP-dependent RNA helicase, to bind the mRNA. Upf1 binds to Upf2, which likely activates the helicase activity of Upf1 (Clerici et al., 2009). The function of this helicase activity is unclear, but may result in the remodeling of the NMD messenger ribonucleoprotein (mRNP), which in turn allows for subsequent cleavage and recycling steps (Franks et al., 2010). Upf2 binds Upf3 (Kadlec et al., 2004), which interacts with the mRNA. This has been shown in mammalian tissue culture to interact with the exon junction complex (Buchwald et al., 2010; Le Hir et al., 2001), which stimulates NMD in mammalian cells (Gehring et al., 2005; Ivanov et al., 2008; Palacios et al., 2004) and plants (Kertész et al., 2006).

In metazoans additional components are involved in NMD and are thought to provide a regulatory role. Smg1 is a phosphoinositide-like kinase (PIKK) that phosphorylates Upf1 (Grimson et al., 2004). Phospho-Upf1 provides a binding substrate

for the proteins Smg5, Smg7, and Smg6 (Okada-Katsuhata et al., 2011). Smg5 and Smg7 in turn recruit phosphatase 2a to the complex to de-phosphorylate Upf1 (Anders et al., 2003; Page et al., 1999). This phosphorylation cycle may function to recycle components of the NMD complex (Isken et al., 2008). Smg6, an endonuclease, cleaves the mRNA near the premature termination codon, resulting in rapid decay of uncapped and unadenylated free ends by the 3'-5' exosome and the 5'-3' exonuclease Xrn1 (Eberle et al., 2009; Gatfield and Izaurralde, 2004; Huntzinger et al., 2008).

Role of NMD in development

Upf1, *Upf2*, and *Upf3* are conserved in all eukaryotes examined and loss of function of any of these genes results in stabilization of NMD target transcripts. *Upf1, Upf2,* and *Upf3* loss of function mutations in yeast (Cui et al., 1995; Leeds et al., 1991; 1992) or *C. elegans* (Cali et al., 1999; Hodgkin et al., 1989) do not affect viability; thus, NMD is not an essential process in these organisms. However, loss of *Upf1* in plants (Arciga-Reyes et al., 2006; Yoine et al., 2006), *Upf1* or *Upf2* in *Drosophila* (Avery et al., 2011; Frizzell et al., 2012; Metzstein and Krasnow, 2006), zebrafish (Wittkopp et al., 2009), and mice (Medghalchi et al., 2001; Thoren et al., 2010; Weischenfeldt et al., 2008), results in lethality; thus, NMD is likely essential for development in these organisms. The third member of this evolutionarily conserved core, *Upf3*, is not essential for viability in *Drosophila* (Avery et al., 2011). In vertebrate lineages, *Upf3* has undergone a duplication. The two paralogs, *Upf3a* and *Upf3b* (*Upf3x*) are not lethal individually (Huang et al., 2011), however, the double knockout has yet to be performed in mice. Humans carrying a hemizygous mutation in *Upf3b* survive to adulthood with a form of

X-linked mental retardation (Tarpey et al., 2007). *Upf3b* knockout mice are viable (Huang et al., 2011). Taken together, this suggests that the paralogs of *Upf3* compensate for defects that may arise in the other gene (Chan et al., 2009); however, whether this occurs *in vivo* is not yet clear (Huang et al., 2011). The lack of redundancy of *Upf3* in *Drosophila* is interesting as there is only one conserved *Upf3* ortholog, yet loss of function of Upf3 is not lethal (Avery et al., 2011). One hypothesis is that *Upf3* may be required for targeting only a subset on NMD transcripts, thus representing a branch of the NMD pathway that is not absolutely required for viability. The variability in lethality associated with loss of NMD genes could result from functional differences in NMD between organisms; different methods of targeting transcripts and different branches of NMD may have arisen throughout evolution. In plants (Kertész et al., 2006), zebrafish (Wittkopp et al., 2009), and mammals (Le Hir et al., 2001), the exon junction complex (EJC) can stimulate NMD transcript targeting; however, while the EJC is present in *Drosophila*, it does not appear to play a role in NMD (Gatfield et al., 2003). *Upf3* interacts with the EJC in mammalian cells, suggesting that the *Upf3*-EJC may provide an important level of regulation for a vertebrate specific NMD pathway (Buchwald et al., 2010).

Smg7 is the only auxiliary NMD gene conserved in plants (Benkovics et al., 2011; Kerényi et al., 2008), whereas *Smg1*, *Smg5*, *Smg6*, and *Smg7* are conserved between *C. elegans* and vertebrates (Aronoff et al., 2001; Chiu et al., 2003; Culbertson and Leeds, 2003; Denning et al., 2001; Serin et al., 2001). *Smg1*, *Smg5* and *Smg6* are conserved in *Drosophila*, but *Smg7* appears to be absent (Gatfield et al., 2003). Based on sequence and structure, *Smg7* may be functionally redundant with *Smg5* (Chiu et al., 2003). While

the biochemical roles of *Smg1*, *5, 6* and *7* are well understood (Kashima et al., 2006; Okada-Katsuhata et al., 2011), their *in vivo* roles have only recently begun to be examined. *Smg1* is required for embryonic viability in mice (McIlwain et al., 2010), supporting the hypothesis that NMD is essential for viability in complex organisms. However, morpholino knockdown of *Smg1* in zebrafish (Wittkopp et al., 2009) and mutant analysis in *Drosophila* (Chen et al., 2005; Metzstein and Krasnow, 2006) showed *Smg1* plays only a minor role in development and NMD function in these organisms, suggesting that phosphorylation of Upf1 is not essential for NMD. However, this interpretation is confounded by experiments in zebrafish, which show morpholino knockdown of *Smg5* or *Smg7* results in severe developmental defects (Wittkopp et al., 2009). This suggests that dephosphorylation of Upf1 is important for viability in zebrafish. Alternatively, Smg5 and Smg7 may be involved in developmental processes other than NMD (Wittkopp et al., 2009). Analysis of NMD genes by morpholinos is subject to off-target effects, which could result in the developmental defects observed. Additionally, the zebrafish morphant animals were not tested directly for stabilization of NMD targets, so the requirement for each of the genes in the NMD process *per se* is not known.

Analysis of NMD function *in vivo*

Several hypotheses have been proposed as to why NMD may cause organismal lethality, two of which are the "vacuum cleaner" and "swiss-army knife" models (Neu-Yilik et al., 2004). The vacuum cleaner model suggests NMD acts to clear cells of excess debris from nonsense transcripts, and lethality in NMD mutants results from a build-up of

defective mRNA and truncated proteins that should have been degraded. The swiss-army knife model suggests NMD clears cells of a small number of specific transcripts, the regulation of which is essential for development (Neu-Yilik et al., 2004). Other models postulate that lethality occurs because NMD genes have other functions unrelated to NMD in more complex organisms (Hwang and Maquat, 2011). The best way to test these models is through methods that allow us to investigate NMD in different cells, tissues, and throughout developmental stages. Development and analysis of allelic series of NMD genes in *Drosophila* is one way to begin differentiating these different models of NMD function *in vivo*.

Our screen for genes involved in NMD in *Drosophila* isolated six alleles of *Smg6*, the first alleles of this gene to be isolated in this organism (Frizzell et al., 2012). Our results indicate that *Smg6* is not absolutely required for viability or for NMD function *in vivo*. Analysis of *Smg6* has been performed by morpholino knockdown in zebrafish (Wittkopp et al., 2009) and suggested loss of *Smg6* was equal to or more severe than loss of *Upf1* developmentally. This surprising result could be explained by morpholino off-target affects. Additionally, mutant analysis in *C. elegans* has shown *Smg6* is involved in NMD (Cali and Anderson, 1998; Hodgkin et al., 1989; Page et al., 1999). However, since the NMD genes are not required for viability in this organism, we cannot ascertain essential developmental roles associated with loss of individual NMD components. Recent studies in *Drosophila* S2 cells (Huntzinger et al., 2008) and mammalian cells (Eberle et al., 2009) demonstrated that *Smg6* is an endonuclease that cleaves NMD target mRNAs near the premature termination codon (Gatfield and Izaurralde, 2004). In *Drosophila*, it was suggested that the endonuclease function of *Smg6* is the primary mode

of cleavage resulting in the rapid decay of target transcripts (Gatfield and Izaurralde, 2004; Huntzinger et al., 2008). Work in yeast (He and Jacobson, 2001) and mammalian cells (Chen and Shyu, 2003; Couttet and Grange, 2004; Muhlrad and Parker, 1994) has shown a deadenylation independent decapping mechanism is also utilized by NMD. The partial loss of NMD function seen in *Drosophila Smg6* mutants suggests an alternative degradation pathway, like the one seen in yeast and mammals, may also be utilized in *Drosophila*. Chapter 4 addresses methods to further examine the role of *Smg6 in vivo*, focusing on the endonuclease domain of *Smg6*, which is disrupted in many of the mutant alleles (Frizzell et al., 2012). Additionally, this chapter addresses ideas to identify other degradation pathways utilized by NMD in *Drosophila*.

How does NMD distinguish between normal or premature termination codons?

Biochemical analysis of NMD genes has resulted in a model of how PTC-containing transcripts are targeted by NMD. Two alternative, though nonexclusive, models have been proposed based on studies in a number of systems. A mammalian cell culture-based model for PTC recognition depends on the exon junction complex (EJC). The EJC-complex model was based on a series of observations that the presence of an intron downstream of the PTC caused destabilization of the transcript (Belgrader et al., 1994; Carter et al., 1996; Cheng et al., 1994; Zhang et al., 1998). The EJC is a multiprotein complex that is deposited on the mRNA upon splicing (Le Hir et al., 2000). The EJC remains associated with the RNA until the pioneer (first) round of translation (Lejeune and Maquat, 2005), at which point the ribosome displaces EJCs from the

mRNA. Since genes typically encode a termination codon in the final exon, all EJCs are displaced by the translating ribosome. This results in no EJC bound to naturally terminating transcripts, allowing for interactions between the ribosome, release factors and poly-A binding protein to stimulate efficient translation termination (Fig. 1.1). However, if a premature stop is recognized before an exon-exon junction, the ribosome will not continue to remove the EJCs downstream of the PTC, leaving any associated downstream EJC as a mark of a PTC. A component of the NMD complex, Upf3, interacts with the EJC upon translocation of the mRNA from the nucleus to the cytoplasm (Buchwald et al., 2010; Singh et al., 2007). The interaction between Upf3, the EJC, NMD components, and the translation release factors eRF1 and eRF3 is called the SURF complex (Kashima et al., 2006) (Fig. 1.2). The presence of an NMD component associated with the mRNA in the SURF complex is thought to prime transcripts for targeting by NMD. However, this model does not account for messages that have no introns, or have PTCs in the final exon that are still subject to NMD. Additionally, experiments in *Drosophila* tissue culture indicate the EJC does not contribute to transcript recognition, suggesting this method of targeting is not conserved (Gatfield et al., 2003).

An alternative model was primarily established in yeast (Amrani et al., 2004), and is now supported by data from *C. elegans*, *Drosophila*, plants and mammalian tissue culture. This model postulates the length of the 3' UTR is key to determining transcript stability (Amrani et al., 2004). The mechanism by which length is measured involves interactions between the translating ribosome and proteins associated with the poly-A-tail (Behm-Ansmant et al., 2007). When a translating ribosome recognizes the termination

codon, it interacts with the release factors eRF1 and eRF3. Efficient termination occurs when release factors also interact with poly-A binding protein (PABP) associated with the 3'UTR. Thus, when the translating ribosome terminates at a normal stop codon, this interaction occurs efficiently, which results in ribosome release and recycling of the transcript for entry into steady-state translation. In the case of a premature termination codon, the release factors cannot interact with poly-A-binding protein due to an increase in the physical distance between the factors. The NMD protein, Upf1, interacts with the release factors in a competitive manner that prevents efficient binding of poly-A-binding protein. Upf1 then remains associated with the mRNA and recruits other NMD proteins to the transcript, which stimulates cleavage and decay (Eberle et al., 2008; Hogg and Goff, 2010; Singh et al., 2008) (Fig. 1.3). Currently, this is thought to be the primary mode of target recognition in all organisms, with the EJC model stimulating NMD in some specific cases in invertebrates (Singh et al., 2008) and plants (Kertész et al., 2006).

Genome wide studies in yeast (He et al., 2003), *C. elegans* (Barberan-Soler et al., 2009; Ramani et al., 2009), *Drosophila* S2 cells (Hansen et al., 2009; Rehwinkel et al., 2005), and mammals (Lewis et al., 2003; Nguyen et al., 2012) have shown NMD targets a variety of transcripts, up to 10% of the genomes in some organisms. However, many of these targets lack the features that are predicted to be required for NMD targeting, including uORFs, long 3'UTRs, or introns in the 3'UTR (Kervestin and Jacobson, 2012). The lack of known features in many NMD targets suggests that other previously unknown features may also be required for targeting by NMD. The fluorescent NMD reporter utilized in our screen uses the simian viral small t antigen intron and polyadenylation signal as a 3'UTR; this SV40 3'UTR is targeted for degradation by

NMD (Metzstein and Krasnow, 2006), but is unlike many reporters used in tissue culture and *C. elegans*, as it lacks a premature termination codon. Chapter 5 discusses work performed to map the cis-acting factors within the SV40 3'UTR required for targeting and degradation by NMD.

Summary

Work described in this thesis addresses three main questions related to the function of the NMD pathway. Chapters 2 and 3 describe work establishing whether all NMD genes have been identified. Chapter 4 addresses how NMD genes can be investigated functionally. Finally, Chapter 5 investigates NMD targeting of the SV40 3'UTR and how new and previously described methods may be utilized for NMD targeting models in *Drosophila*.

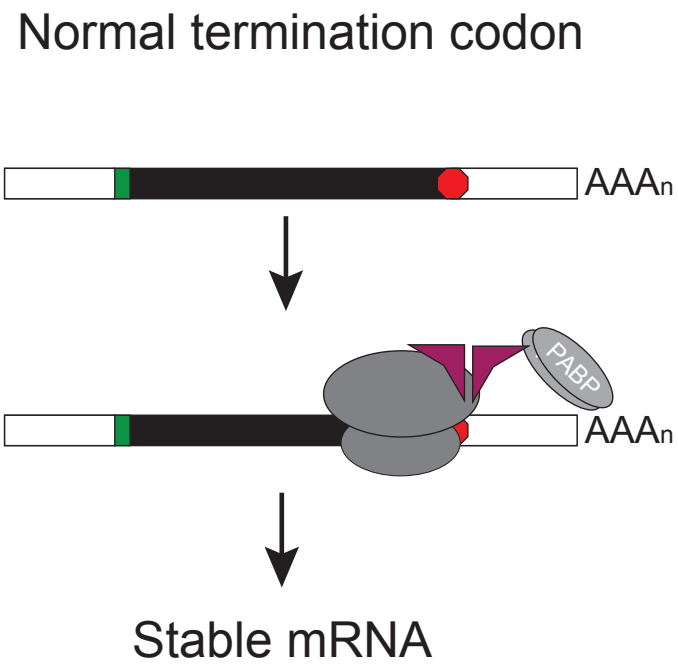


Figure 1.1 Normal translation termination. mRNA containing a normal termination codon is translated by the ribosome; release factors bind the ribosome upon recognition of the termination codon, and interact with poly-A binding protein associated with the poly-A tail. This signals efficient translation termination and the mRNA enters steady state translation with a low rate of decay. Ribosome is represented in dark grey. Release factors are represented by magenta triangles.

Figure 1.2 PTC-containing mRNA with EJC bound. mRNA containing a PTC is translated by the ribosome, the release factors bind the ribosome upon recognition of the termination codon, however recognition of the PTC-bound mRNA is stimulated by the presence of the EJC. The NMD factor Upf3 binds the EJC and recruits other NMD components to the PTC-containing transcript, resulting in cleavage and degradation of the transcript. Release factors are represented by magenta triangles. Ribosome is represented in dark grey. Release factors are represented by magenta triangles.

Figure 1.2 Continued

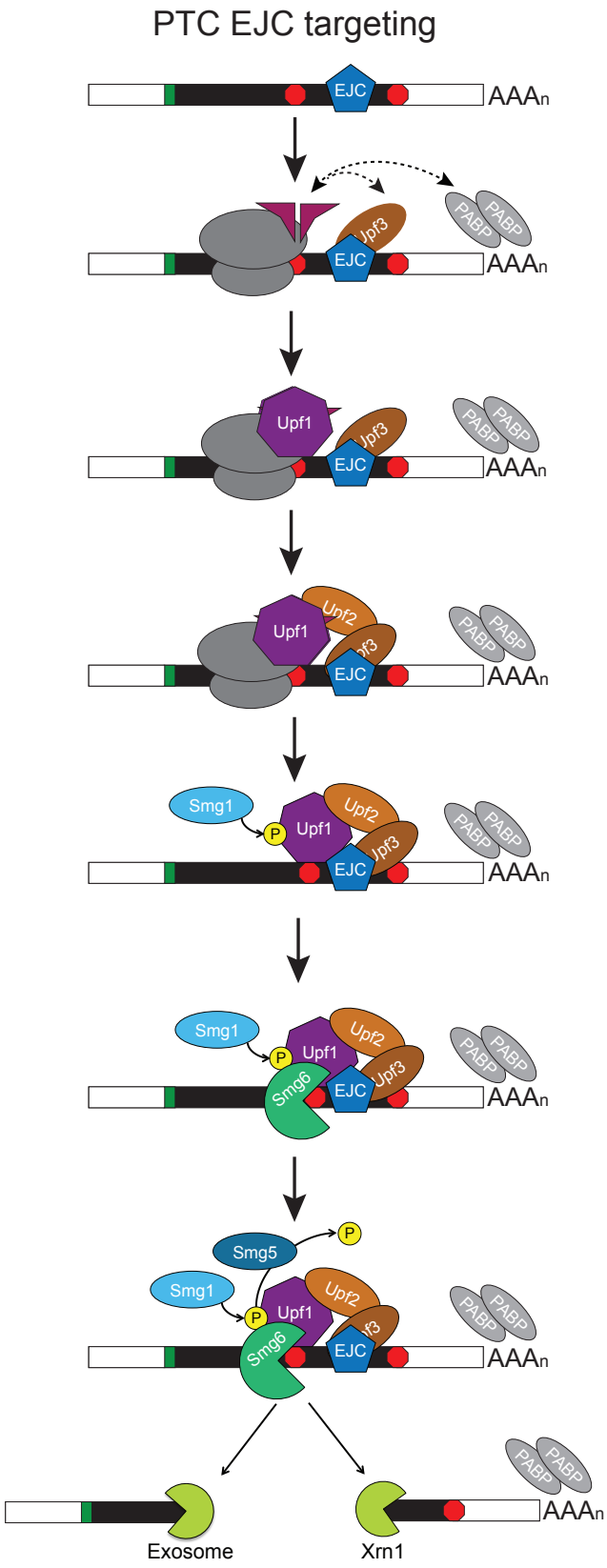
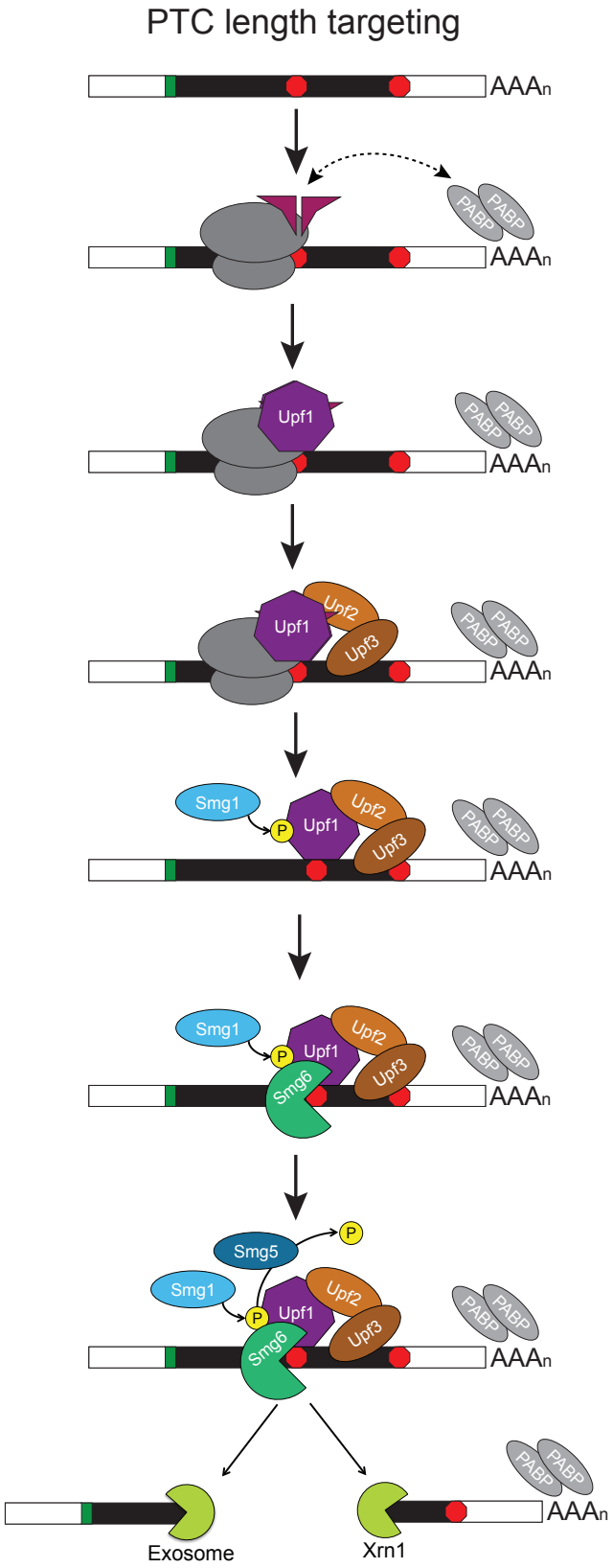


Figure 1.3 Targeting for PTC-containing mRNA, no EJC. mRNA containing a PTC is translated by the ribosome, the release factors bind the ribosome upon recognition of the termination codon, however the distance between the release factors and the poly-A tail prohibits efficient translation termination. Upf1 competitively interacts with poly-A binding protein to bind release factors, which stimulates assembly of the rest of the NMD factors, resulting in cleavage of the PTC-containing transcript and rapid decay. Ribosome is represented in dark grey. Release factors are represented by magenta triangles.

Figure 1.3 Continued



References

Amrani, N., Ganesan, R., Kervestin, S., Mangus, D.A., Ghosh, S., and Jacobson, A. (2004). A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature *432*, 112–118.

Anastasaki, C., Longman, D., Capper, A., Patton, E.E., and Cáceres, J.F. (2011). Dhx34 and Nbas function in the NMD pathway and are required for embryonic development in zebrafish. Nucl. Acids Res. *39*, 3686–3694.

Anders, K.R., Grimson, A., and Anderson, P. (2003). SMG-5, required for C.elegans nonsense-mediated mRNA decay, associates with SMG-2 and protein phosphatase 2A. Embo J. *22*, 641–650.

Arciga-Reyes, L., Wootton, L., Kieffer, M., and Davies, B. (2006). UPF1 is required for nonsense-mediated mRNA decay (NMD) and RNAi in Arabidopsis. Plant J. *47*, 480–489.

Aronoff, R., Baran, R., and Hodgkin, J. (2001). Molecular identification of smg-4, required for mRNA surveillance in C. elegans. Gene *268*, 153–164.

Avery, P., Vicente-Crespo, M., Francis, D., Nashchekina, O., Alonso, C.R., and Palacios, I.M. (2011). Drosophila Upf1 and Upf2 loss of function inhibits cell growth and causes animal death in a Upf3-independent manner. Rna *17*, 624–638.

Barberan-Soler, S., Lambert, N.J., and Zahler, A.M. (2009). Global analysis of alternative splicing uncovers developmental regulation of nonsense-mediated decay in C. elegans. Rna *15*, 1652–1660.

Baserga, S.J., and Benz, E.J. (1988). Nonsense mutations in the human beta-globin gene affect mRNA metabolism. Proc. Natl. Acad. Sci. USA *85*, 2056–2060.

Behm-Ansmant, I., Gatfield, D., Rehwinkel, J., Hilgers, V., and Izaurralde, E. (2007). A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. Embo J. *26*, 1591–1601.

Belgrader, P., Cheng, J., Zhou, X., Stephenson, L.S., and Maquat, L.E. (1994). Mammalian nonsense codons can be cis effectors of nuclear mRNA half-life. Mol. Cell. Biol. *14*, 8219–8228.

Benkovics, A.H., Nyikó, T., Mérai, Z., Silhavy, D., and Bisztray, G.D. (2011). Functional analysis of the grapevine paralogs of the SMG7 NMD factor using a heterolog VIGS-based gene depletion-complementation system. Plant Mol. Biol. *75*, 277–290.

Besse, F., and Ephrussi, A. (2008). Translational control of localized mRNAs: restricting protein synthesis in space and time. Nat. Rev. Mol. Cell Biol. *9*, 971–980.

Buchwald, G., Ebert, J., Basquin, C., Sauliere, J., Jayachandran, U., Bono, F., Le Hir, H., and Conti, E. (2010). Insights into the recruitment of the NMD machinery from the

crystal structure of a core EJC-UPF3b complex. Proc. Natl. Acad. Sci. USA *107*, 10050–10055.

Cali, B.M., and Anderson, P. (1998). mRNA surveillance mitigates genetic dominance in Caenorhabditis elegans. Mol. Gen. Genet. *260*, 176–184.

Cali, B.M., Kuchma, S.L., Latham, J., and Anderson, P. (1999). smg-7 is required for mRNA surveillance in Caenorhabditis elegans. Genetics *151*, 605–616.

Carter, M.S., Li, S., and Wilkinson, M.F. (1996). A splicing-dependent regulatory mechanism that detects translation signals. Embo J. *15*, 5965–5975.

Chan, W.-K., Bhalla, A.D., Le Hir, H., Nguyen, L.S., Huang, L., Gécz, J., and Wilkinson, M.F. (2009). A UPF3-mediated regulatory switch that maintains RNA surveillance. Nat. Struct. Mol. Biol. *16*, 747–753.

Chen, C.-Y.A., and Shyu, A.-B. (2003). Rapid deadenylation triggered by a nonsense codon precedes decay of the RNA body in a mammalian cytoplasmic nonsense-mediated decay pathway. Mol. Cell. Biol. *23*, 4805–4813.

Chen, Z., Smith, K.R., Batterham, P., and Robin, C. (2005). Smg1 nonsense mutations do not abolish nonsense-mediated mRNA decay in Drosophila melanogaster. Genetics *171*, 403–406.

Cheng, J., Belgrader, P., Zhou, X., and Maquat, L.E. (1994). Introns are cis effectors of the nonsense-codon-mediated reduction in nuclear mRNA abundance. Mol. Cell. Biol. *14*, 6317–6325.

Chia, W., Savakis, C., Karp, R., and Ashburner, M. (1987). Adhn4 of Drosophila melanogaster is a nonsense mutation. Nucl. Acids Res. *15*, 3931.

Chiu, S.-Y., Serin, G., Ohara, O., and Maquat, L.E. (2003). Characterization of human Smg5/7a: a protein with similarities to Caenorhabditis elegans SMG5 and SMG7 that functions in the dephosphorylation of Upf1. Rna *9*, 77–87.

Clerici, M., Mourão, A., Gutsche, I., Gehring, N.H., Hentze, M.W., Kulozik, A.E., Kadlec, J., Sattler, M., and Cusack, S. (2009). Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2. Embo J. *28*, 2293–2306.

Couttet, P., and Grange, T. (2004). Premature termination codons enhance mRNA decapping in human cells. Nucl. Acids Res. *32*, 488–494.

Cui, Y., Hagan, K.W., Zhang, S., and Peltz, S.W. (1995). Identification and characterization of genes that are required for the accelerated degradation of mRNAs containing a premature translational termination codon. Genes Dev. *9*, 423–436.

Culbertson, M.R., Underbrink, K.M., and Fink, G.R. (1980). Frameshift suppression Saccharomyces cerevisiae. II. Genetic properties of group II suppressors. Genetics *95*,

833–853.

Culbertson, M.R., and Leeds, P.F. (2003). Looking at mRNA decay pathways through the window of molecular evolution. Curr. Opin. Genetics Dev. *13*, 207–214.

Cummins, C.M., Gaber, R.F., Culbertson, M.R., Mann, R., and Fink, G.R. (1980). Frameshift suppression in Saccharomyces cerevisiae. III. Isolation and genetic properties of group III suppressors. Genetics *95*, 855–879.

Denning, G., Jamieson, L., Maquat, L.E., Thompson, E.A., and Fields, A.P. (2001). Cloning of a novel phosphatidylinositol kinase-related kinase: characterization of the human SMG-1 RNA surveillance protein. J. Biol. Chem. *276*, 22709–22714.

Donahue, T.F., Farabaugh, P.J., and Fink, G.R. (1981). Suppressible four-base glycine and proline codons in yeast. Science *212*, 455–457.

Eberle, A.B., Lykke-Andersen, S., Mühlemann, O., and Jensen, T.H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat. Struct. Mol. Biol. *16*, 49–55.

Eberle, A.B., Stalder, L., Mathys, H., Zamudio Orozco, R., and Mühlemann, O. (2008). Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. PLoS Biol. *6*, e92.

Franks, T.M., Singh, G., and Lykke-Andersen, J. (2010). Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense-mediated mRNA decay. Cell *143*, 938–950.

Frizzell, K.A., Rynearson, S.G., and Metzstein, M.M. (2012). Drosophila mutants show NMD pathway activity is reduced, but not eliminated, in the absence of Smg6. Rna *18*, 1475–1486.

Fuller-Pace, F.V. (2006). DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation. Nucl. Acids Res. *34*, 4206–4215.

Gatfield, D., and Izaurralde, E. (2004). Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. Nature *429*, 575–578.

Gatfield, D., Unterholzner, L., Ciccarelli, F.D., Bork, P., and Izaurralde, E. (2003). Nonsense-mediated mRNA decay in Drosophila: at the intersection of the yeast and mammalian pathways. Embo J. *22*, 3960–3970.

Gehring, N.H., Kunz, J.B., Neu-Yilik, G., Breit, S., Viegas, M.H., Hentze, M.W., and Kulozik, A.E. (2005). Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. Mol. Cell *20*, 65–75.

Grimson, A., O'Connor, S., Newman, C.L., and Anderson, P. (2004). SMG-1 is a

phosphatidylinositol kinase-related protein kinase required for nonsense-mediated mRNA decay in Caenorhabditis elegans. Mol. Cell. Biol. *24*, 7483–7490.

Hansen, K.D., Lareau, L.F., Blanchette, M., Green, R.E., Meng, Q., Rehwinkel, J., Gallusser, F.L., Izaurralde, E., Rio, D.C., Dudoit, S., et al. (2009). Genome-wide identification of alternative splice forms down-regulated by nonsense-mediated mRNA decay in Drosophila. PLoS Genet. *5*, e1000525.

Harigaya, Y., and Parker, R. (2010). No-go decay: a quality control mechanism for RNA in translation. Wiley Interdiscip. Rev. RNA *1*, 132–141.

He, F., and Jacobson, A. (1995). Identification of a novel component of the nonsense-mediated mRNA decay pathway by use of an interacting protein screen. Genes Dev. *9*, 437–454.

He, F., and Jacobson, A. (2001). Upf1p, Nmd2p, and Upf3p regulate the decapping and exonucleolytic degradation of both nonsense-containing mRNAs and wild-type mRNAs. Mol. Cell. Biol. *21*, 1515–1530.

He, F., Brown, A.H., and Jacobson, A. (1997). Upf1p, Nmd2p, and Upf3p are interacting components of the yeast nonsense-mediated mRNA decay pathway. Mol. Cell. Biol. *17*, 1580–1594.

He, F., Li, X., Spatrick, P., Casillo, R., Dong, S., and Jacobson, A. (2003). Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5“ to 3” mRNA decay pathways in yeast. Mol. Cell *12*, 1439–1452.

Hodgkin, J., Papp, A., Pulak, R., Ambros, V., and Anderson, P. (1989). A new kind of informational suppression in the nematode Caenorhabditis elegans. Genetics *123*, 301–313.

Hogg, J.R., and Goff, S.P. (2010). Upf1 Senses 3'UTR Length to Potentiate mRNA Decay. Cell *143*, 379–389.

Huang, L., Lou, C.-H., Chan, W.-K., Shum, E.Y., Shao, A., Stone, E., Karam, R., Song, H.-W., and Wilkinson, M.F. (2011). RNA Homeostasis Governed by Cell Type-Specific and Branched Feedback Loops Acting on NMD. Mol. Cell *43*, 950–961.

Huntzinger, E., Kashima, I., Fauser, M., Saulière, J., and Izaurralde, E. (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. Rna *14*, 2609–2617.

Hwang, J., and Maquat, L.E. (2011). Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: to die or not to die, that is the question. Curr. Opin. Genetics Dev. *21*, 422–430.

Isken, O., Kim, Y.K., Hosoda, N., Mayeur, G.L., Hershey, J.W.B., and Maquat, L.E. (2008). Upf1 phosphorylation triggers translational repression during nonsense-mediated

mRNA decay. Cell *133*, 314–327.

Ivanov, P.V., Gehring, N.H., Kunz, J.B., Hentze, M.W., and Kulozik, A.E. (2008). Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. Embo J. *27*, 736–747.

Kadlec, J., Izaurralde, E., and Cusack, S. (2004). The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3. Nat. Struct. Mol. Biol. *11*, 330–337.

Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G., and Ohno, S. (2006). Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. Genes Dev. *20*, 355–367.

Kerényi, Z., Mérai, Z., Hiripi, L., Benkovics, A.H., Gyula, P., Lacomme, C., Barta, E., Nagy, F., and Silhavy, D. (2008). Inter-kingdom conservation of mechanism of nonsense-mediated mRNA decay. Embo J. *27*, 1585–1595.

Kertész, S., Kerényi, Z., Mérai, Z., Bartos, I., Pálfy, T., Barta, E., and Silhavy, D. (2006). Both introns and long 3'-UTRs operate as cis-acting elements to trigger nonsense-mediated decay in plants. Nucl. Acids Res. *34*, 6147–6157.

Kervestin, S., and Jacobson, A. (2012). NMD: a multifaceted response to premature translational termination. Nat. Rev. Mol. Cell Biol. *13*, 700–712.

Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M.J. (2001). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. Embo J. *20*, 4987–4997.

Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. (2000). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. Embo J. *19*, 6860–6869.

Lee, B.S., and Culbertson, M.R. (1995). Identification of an additional gene required for eukaryotic nonsense mRNA turnover. Proc. Natl. Acad. Sci. USA *92*, 10354–10358.

Lee, T., and Luo, L. (2001). Mosaic analysis with a repressible cell marker (MARCM) for Drosophila neural development. Trends Neurosci. *24*, 251–254.

Leeds, P., Peltz, S.W., Jacobson, A., and Culbertson, M.R. (1991). The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon. Genes Dev. *5*, 2303–2314.

Leeds, P., Wood, J.M., Lee, B.S., and Culbertson, M.R. (1992). Gene products that promote mRNA turnover in Saccharomyces cerevisiae. Mol. Cell. Biol. *12*, 2165–2177.

Lejeune, F., and Maquat, L.E. (2005). Mechanistic links between nonsense-mediated

mRNA decay and pre-mRNA splicing in mammalian cells. Curr. Opin. Cell Biol. *17*, 309–315.

Lewis, B.P., Green, R.E., and Brenner, S.E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc. Natl. Acad. Sci. USA *100*, 189–192.

Longman, D., Plasterk, R.H.A., Johnstone, I.L., and Cáceres, J.F. (2007). Mechanistic insights and identification of two novel factors in the C. elegans NMD pathway. Genes Dev. *21*, 1075–1085.

Losson, R., and Lacroute, F. (1979). Interference of nonsense mutations with eukaryotic messenger RNA stability. Proc. Natl. Acad. Sci. USA *76*, 5134–5137.

McGlincy, N.J., and Smith, C.W.J. (2008). Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? Trends Biochem. Sci. *33*, 385–393.

McIlwain, D.R., Pan, Q., Reilly, P.T., Elia, A.J., McCracken, S., Wakeham, A.C., Itie-Youten, A., Blencowe, B.J., and Mak, T.W. (2010). Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. Proc. Natl. Acad. Sci. USA *107*, 12186–12191.

Medghalchi, S.M., Frischmeyer, P.A., Mendell, J.T., Kelly, A.G., Lawler, A.M., and Dietz, H.C. (2001). Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. Hum. Mol. Genet. *10*, 99–105.

Metzstein, M.M., and Krasnow, M.A. (2006). Functions of the nonsense-mediated mRNA decay pathway in Drosophila development. PLoS Genet. *2*, e180.

Muhlrad, D., and Parker, R. (1994). Premature translational termination triggers mRNA decapping. Nature *370*, 578–581.

Mühlemann, O., Eberle, A.B., Stalder, L., and Zamudio Orozco, R. (2008). Recognition and elimination of nonsense mRNA. Biochim. Biophys. Acta. *1779*, 538–549.

Neu-Yilik, G., Gehring, N.H., Hentze, M.W., and Kulozik, A.E. (2004). Nonsense-mediated mRNA decay: from vacuum cleaner to Swiss army knife. Genome Biol. *5*, 218.

Nguyen, L.S., Jolly, L., Shoubridge, C., Chan, W.K., Huang, L., Laumonnier, F., Raynaud, M., Hackett, A., Field, M., Rodriguez, J., et al. (2012). Transcriptome profiling of UPF3B/NMD-deficient lymphoblastoid cells from patients with various forms of intellectual disability. Mol. Psychiatry *17*, 1103–1115.

Nicholson, P., Yepiskoposyan, H., Metze, S., Zamudio Orozco, R., Kleinschmidt, N., and Mühlemann, O. (2009). Nonsense-mediated mRNA decay in human cells: mechanistic insights, functions beyond quality control and the double-life of NMD factors. Cell Mol. Life. Sci. *67*, 677–700.

Okada-Katsuhata, Y., Yamashita, A., Kutsuzawa, K., Izumi, N., Hirahara, F., and Ohno, S. (2011). N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. Nucl. Acids Res. 1–16.

Okamura, K. (2012). Diversity of animal small RNA pathways and their biological utility. Wiley Interdiscip. Rev. RNA *3*, 351–368.

Page, M.F., Carr, B., Anders, K.R., Grimson, A., and Anderson, P. (1999). SMG-2 is a phosphorylated protein required for mRNA surveillance in Caenorhabditis elegans and related to Upf1p of yeast. Mol. Cell. Biol. *19*, 5943–5951.

Palacios, I.M., Gatfield, D., St Johnston, D., and Izaurralde, E. (2004). An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. Nature *427*, 753–757.

Peltz, S.W., Brown, A.H., and Jacobson, A. (1993). mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor. Genes Dev. *7*, 1737–1754.

Qiu, S., Adema, C., and Lane, T. (2005). A computational study of off-target effects of RNA interference. Nucl. Acids Res. *33*, 1834–1847.

Qu, W., Ren, C., Li, Y., Shi, J., Zhang, J., Wang, X., Hang, X., Lu, Y., Zhao, D., and Zhang, C. (2011). Reliability analysis of the Ahringer Caenorhabditis elegans RNAi feeding library: a guide for genome-wide screens. BMC Genomics *12*, 170.

Ramani, A.K., Nelson, A.C., Kapranov, P., Bell, I., Gingeras, T.R., and Fraser, A.G. (2009). High resolution transcriptome maps for wild-type and nonsense-mediated decay-defective Caenorhabditis elegans. Genome Biol. *10*, R101.

Rehwinkel, J., Letunic, I., Raes, J., Bork, P., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. Rna *11*, 1530–1544.

Rosains, J., and Mango, S.E. (2012). Genetic characterization of smg-8 mutants reveals no role in C. elegans nonsense mediated decay. PLoS ONE *7*, e49490.

Serin, G., Gersappe, A., Black, J.D., Aronoff, R., and Maquat, L.E. (2001). Identification and characterization of human orthologues to Saccharomyces cerevisiae Upf2 protein and Upf3 protein (Caenorhabditis elegans SMG-4). Mol. Cell. Biol. *21*, 209–223.

Singh, G., Jakob, S., Kleedehn, M.G., and Lykke-Andersen, J. (2007). Communication with the exon-junction complex and activation of nonsense-mediated decay by human Upf proteins occur in the cytoplasm. Mol. Cell *27*, 780–792.

Singh, G., Rebbapragada, I., and Lykke-Andersen, J. (2008). A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. PLoS Biol. *6*, e111.

Tarpey, P.S., Raymond, F.L., Nguyen, L.S., Rodriguez, J., Hackett, A., Vandeleur, L., Smith, R., Shoubridge, C., Edkins, S., Stevens, C., et al. (2007). Mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation. Nat. Genet. *39*, 1127–1133.

Thoren, L.A., Nørgaard, G.A., Weischenfeldt, J., Waage, J., Jakobsen, J.S., Damgaard, I., Bergström, F.C., Blom, A.M., Borup, R., Bisgaard, H.C., et al. (2010). UPF2 is a critical regulator of liver development, function and regeneration. PLoS ONE *5*, e11650.

Timmons, L., Court, D.L., and Fire, A. (2001). Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans. Gene *263*, 103–112.

Vasudevan, S., Peltz, S.W., and Wilusz, C.J. (2002). Non-stop decay--a new mRNA surveillance pathway. Bioessays *24*, 785–788.

Weischenfeldt, J., Damgaard, I., Bryder, D., Theilgaard-Mönch, K., Thoren, L.A., Nielsen, F.C., Jacobsen, S.E.W., Nerlov, C., and Porse, B.T. (2008). NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. Genes Dev. *22*, 1381–1396.

Wittkopp, N., Huntzinger, E., Weiler, C., Sauliere, J., Schmidt, S., Sonawane, M., and Izaurralde, E. (2009). Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic development and survival. Mol. Cell. Biol. *29*, 3517–3528.

Xu, T., and Rubin, G.M. (1993). Analysis of genetic mosaics in developing and adult Drosophila tissues. Development *117*, 1223–1237.

Yamashita, A., Izumi, N., Kashima, I., Ohnishi, T., Saari, B., Katsuhata, Y., Muramatsu, R., Morita, T., Iwamatsu, A., Hachiya, T., et al. (2009). SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. Genes Dev. *23*, 1091–1105.

Yoine, M., Nishii, T., and Nakamura, K. (2006). Arabidopsis UPF1 RNA helicase for nonsense-mediated mRNA decay is involved in seed size control and is essential for growth. Plant Cell Phys. *47*, 572–580.

Zhang, J., Sun, X., Qian, Y., LaDuca, J.P., and Maquat, L.E. (1998). At least one intron is required for the nonsense-mediated decay of triosephosphate isomerase mRNA: a possible link between nuclear splicing and cytoplasmic translation. Mol. Cell. Biol. *18*, 5272–5283.

CHAPTER 2

DROSOPHILA MUTANTS SHOW NMD PATHWAY ACTIVITY IS REDUCED, BUT NOT ELIMINATED, IN THE ABSENCE OF *SMG6*

Reprint of: Frizzell, K. A., Rynearson, S. G., and Metzstein, M. M. (2012) *Drosophila* mutants show NMD pathway activity is reduced, but not eliminated, in the absence of *Smg6*. RNA *18*, 8.

REPORT

Drosophila mutants show NMD pathway activity is reduced, but not eliminated, in the absence of *Smg6*

KIMBERLY A. FRIZZELL, SHAWN G. RYNEARSON, and MARK M. METZSTEIN[1]
Department of Human Genetics, University of Utah, Salt Lake City, Utah 84112, USA

ABSTRACT

The nonsense-mediated mRNA decay (NMD) pathway is best known for targeting mutant mRNAs containing premature termination codons for rapid degradation, but it is also required for regulation of many endogenous transcripts. Components of the NMD pathway were originally identified by forward genetic screens in yeast and *Caenorhabditis elegans*. In other organisms, the NMD pathway has been investigated by studying the homologs of these genes. We present here the first unbiased genetic screen in *Drosophila* designed specifically to identify genes involved in NMD. By using a highly efficient genetic mosaic approach, we have screened ~40% of the *Drosophila* genome and isolated more than 40 alleles of genes required for NMD. We focus on alleles we have obtained in two known NMD components: *Upf2* and *Smg6*. Our analysis of multiple alleles of the core NMD component *Upf2* reveals that the Upf2 requirement in NMD may be separate from its requirement for viability, indicating additional critical cellular roles for this protein. Our alleles of *Smg6* are the first point mutations obtained in *Drosophila*, and we find that *Smg6* has both endonucleolytic and nonendonucleolytic roles in NMD. Thus, our genetic screens have revealed that *Drosophila* NMD factors play distinct roles in target regulation, similar to what is found in mammals, but distinct from the relatively similar requirements for NMD genes observed in *C. elegans* and yeast.

Keywords: *Drosophila*; NMD; *Smg6*; *Upf2*; nonsense-mediated decay

INTRODUCTION

Nonsense-mediated mRNA decay (NMD) is a cellular pathway that selectively targets mRNAs containing premature termination codons (PTCs) for rapid degradation (Chang et al. 2007) and is thought to function to prevent the translation and build-up of potentially harmful truncated protein products (Cali and Anderson 1998). *Trans*-acting factors involved in NMD were originally identified in suppressor screens performed in yeast (Leeds et al. 1991, 1992) and *Caenorhabditis elegans* (Hodgkin et al. 1989; Cali et al. 1999). These screens took advantage of PTC-containing alleles of genes whose truncated protein products, rather than being toxic, possess residual function. Since NMD targets these transcripts, the truncated proteins cannot be made, effectively resulting in a complete loss-of-function phenotype. Loss of function of genes required for NMD results in stabilization of such PTC-containing transcripts, allowing for production of the truncated protein products and amelioration of the mutant phenotype. These screens identified three genes required for NMD in yeast—*UPF1*, *UPF2* (also called NMD2), and *UPF3* (*Upf, up-frameshift suppressor*) (Leeds et al. 1991, 1992)—and seven in *C. elegans*—*smg-1* through *smg-7* (*smg, suppressor with morphogenetic effect on genitalia*) (Hodgkin et al. 1989; Cali et al. 1999). Cloning of these genes has revealed that *UPF1*, *UPF2*, and *UPF3* are orthologous to *smg-2*, *smg-3*, and *smg-4*, respectively, thus defining an evolutionarily conserved core machinery for NMD (Cui et al. 1995; He and Jacobson 1995; Lee and Culbertson 1995; Cali et al. 1999; Page et al. 1999; Aronoff et al. 2001; Serin et al. 2001; Chiu et al. 2003; Grimson et al. 2004). All eukaryotes analyzed contain orthologs of these core genes, and, in all cases examined, they are known to be required for NMD (Kerényi et al. 2008). The four other genes identified in *C. elegans* (*smg-1*, *smg-5*, *smg-6*, *smg-7*) are generally conserved only in metazoa, although there is evidence that a homolog of *Smg7* is also required for NMD in plants (Luke et al. 2007). Smg5 and Smg7 are structurally similar (Fukuhara et al. 2005), although not redundant for their function in NMD, and some organisms, including *Drosophila*, appear only to possess *Smg5* (Gatfield et al. 2003).

[1]**Corresponding author**
E-mail markm@genetics.utah.edu
Article published online ahead of print. Article and publication date are at http://www.rnajournal.org/cgi/doi/10.1261/rna.032821.112.

The design of the suppressor screens used to identify NMD genes demanded that compromising NMD function would not in itself result in lethality. Indeed, this is the case in yeast and *C. elegans*: Complete absence of core NMD components and NMD activity is compatible with viability (Leeds et al. 1991, 1992; Cui et al. 1995; He and Jacobson 1995; Lee and Culbertson 1995; Cali et al. 1999; Page et al. 1999; Aronoff et al. 2001; Grimson et al. 2004). However, for most eukaryotes, disruption of NMD gene function is lethal (Hwang and Maquat 2011), and this essential role has limited the ability to identify NMD genes as suppressors of PTC-containing alleles in these organisms.

Since the initial screens, the only unbiased forward genetic screen that has been performed to identify genes required for NMD has been an RNAi-based approach in *C. elegans*. This screen did not rely on suppression, but instead used an NMD-sensitive reporter whose expression increased upon disruption of NMD activity (Longman et al. 2007). This screen identified two new NMD genes, *smgl-1* and *smgl-2* (*smg lethal*), whose knockdown leads to defective NMD and also organismal lethality. Knockdown of *smgl-1* and *smgl-2* in human cells (Longman et al. 2007) and zebrafish (Anastasaki et al. 2011) also leads to impairment of NMD. However, no mutations in *smgl-1* or *smgl-2* have been reported in any organism.

The primary mechanism that distinguishes premature from normal termination codons seems to involve detection of 3′-UTR length (Amrani et al. 2004; Longman et al. 2007; Eberle et al. 2008). Such a mechanism also plays a role in mammalian cells and plant cells, but there is also an important contribution from interactions between NMD components and splicing-deposited exon junction complexes (Lejeune and Maquat 2005; Kertész et al. 2006; Eberle et al. 2008). Recognition of a PTC is thought to result in the recruitment of the core NMD protein Upf1 (SMG-2 in *C. elegans*), an ATP-dependent RNA-helicase, to the targeted mRNA, triggering degradation (Kashima et al. 2006). In both yeast and mammalian cells, degradation is known to occur through deadenylation-independent decapping, dependent on the recruitment of the decapping proteins DCP1 and DCP2 by Upf1 (Lejeune et al. 2003; Swisher and Parker 2011). This is followed by degradation by 5′–3′ exonucleases, such as XRN1 (He et al. 2003). However, there is now considerable evidence that in metazoan cells, degradation can also occur through endonucleolytic cleavage in the vicinity of the PTC (Gatfield and Izaurralde 2004; Huntzinger et al. 2008; Eberle et al. 2009). The endonuclease responsible for this cleavage is likely to be the NMD component Smg6. Smg6 contains a C-terminally located PIN-endonuclease domain that is required for cleavage of target mRNAs (Glavan et al. 2006; Huntzinger et al. 2008; Eberle et al. 2009). Smg6 is thought to be recruited to target mRNAs by virtue of its 14-3-3-like domain, which can bind phosphorylated Upf1 (Fukuhara et al. 2005). In mammalian cells, Smg6 is also recruited to target mRNAs through direct recruitment to the EJC (Kashima et al. 2010). Smg6 function has primarily been analyzed in cell culture, and while some genetic analysis has been performed in *C. elegans* (Hodgkin et al. 1989; Page et al. 1999), a role for SMG-6 endonuclease function has not yet been analyzed in this organism.

Roles of *Drosophila* NMD factors have been well established in cell culture experiments (Gatfield et al. 2003; Rehwinkel et al. 2005; Huntzinger et al. 2008). These experiments have shown that all the *Drosophila* homologs of known NMD genes are required for NMD and that loss of any of these genes leads to similar defects in NMD. Less is known about how NMD genes function in intact animals. Previous genetic analysis has identified alleles of *Upf1-3* and *Smg1* (Chen et al. 2005; Metzstein and Krasnow 2006; Avery et al. 2011). All of these have been shown to have function in NMD, although, interestingly, some of these roles differ. For instance, while *Upf1* and *Upf2* seem to be absolutely required for NMD and organismal viability, loss of *Smg1* or *Upf3* leads to viable animals with a reduction, but not a complete loss, of NMD pathway activity. Why intact animals and cell culture show these differences is not understood, but probably indicate additional layers of pathway regulation.

Here, we present a novel forward genetic screen in *Drosophila* that uses an NMD-sensitive reporter combined with a genetic mosaic approach. This has allowed us to isolate more than 30 mutations in the known NMD genes *Upf1*, *Upf2*, *Smg1*, and *Smg6*. Our alleles of *Upf2* have revealed that differential interactions of Upf2 with other core NMD components are correlated with whether these alleles are required for viability. The *Smg6* mutations are the first obtained in *Drosophila*, and we have used genetic and molecular criteria to show that several of these are complete loss of function. We find that mutants carrying these null alleles still show considerable NMD pathway function, thus revealing that Smg6 is not absolutely required for NMD in vivo.

RESULTS AND DISCUSSION

Tissue-directed mosaic screen for mutations affecting NMD

To identify novel genes required for NMD as well as alleles of known NMD genes, we performed a large-scale forward genetic screen in *Drosophila*. The screen uses the simian virus 40 small t antigen intron and polyadenylation signal (SV40 3′ UTR), which is targeted for degradation by the NMD pathway in *Drosophila* (Metzstein and Krasnow 2006). Transgenic reporter constructs that use the SV40 3′ UTR as their termination signal result in a low level of expression in wild-type backgrounds; however, the signal increases dramatically when genes involved in NMD are inactivated (Fig. 1A,B). Fluorescent protein reporters with

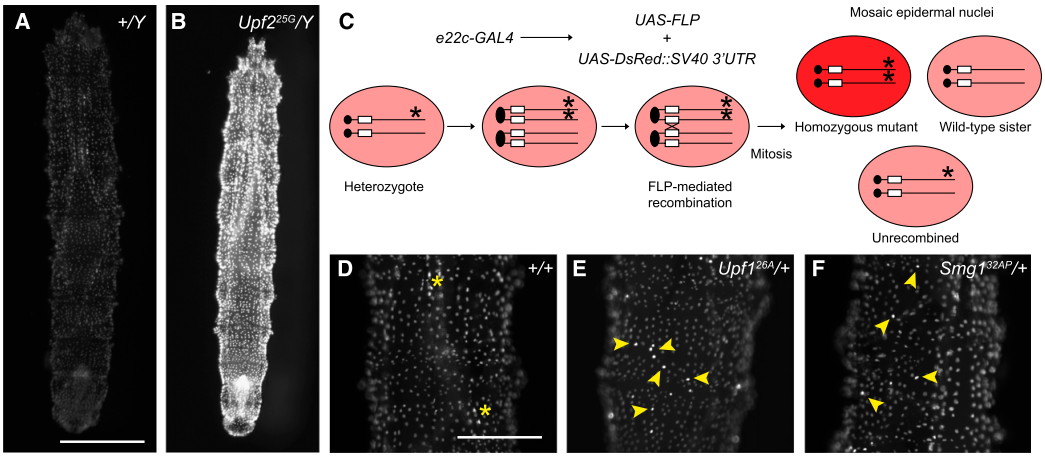


FIGURE 1. Mosaic approach used to detect animals with mutations in genes affecting NMD. (*A*) *y w FRT^{19A}/Y; e22c-GAL4 UAS-nlsDsRed2::SV40 3′UTR/+* L3 (nonmosaic) larva shows uniform, low-level DsRed expression in epidermal nuclei. (*B*) *y w Upf2^{25G} FRT^{19A}/Y; e22c-GAL4, UAS-nlsDsRed2::SV40 3′UTR/+* (nonmosaic) L3 larva, taken at the same exposure as *A*, shows uniform, high-level DsRed expression in epidermal nuclei, due to stabilization of the reporter mRNA. (*C*) *UAS-FLP* expression in the early embryonic ectoderm driven by *e22c-GAL4* catalyzes recombination in heterozygous mitotic cells, resulting in homozygous mutant and wild-type sister cells. Cells that have not undergone recombination remain heterozygous. Expression of an NMD-sensitive reporter *UAS-nlsDsRed2::SV40 3′UTR*, which encodes a nuclear-localized red-fluorescent protein, is driven by *e22c-GAL4* in the larval epidermis. In heterozygous and the wild-type sister cells, the reporter mRNA is targeted for degradation by NMD and shows a low level of DsRed fluorescence (light red). When a gene required for NMD is mutated, the homozygous mutant cell is recognized by an increase in fluorescence (increased red color). (Circles) Centromeres; (rectangles) FRT sites; (*) EMS-generated mutations. (*D*) *y w FRT^{19A}; e22c-GAL4, UAS-nlsDsRed2::SV40 3′UTR, UAS-FLP/+* mosaic L3 larva shows uniform DsRed expression in epidermal nuclei. (Yellow asterisks) Underlying tracheal nuclei that also express DsRed. These are easily distinguished from epidermal cells under the microscope. (*E*) *y w Upf1^{26A} FRT^{19A}/y w FRT^{19A}; e22c-GAL4, UAS-nlsDsRed2::SV40 3′UTR, UAS-FLP/+* L3 larva shows mosaic increased DsRed expression in epidermal nuclei: Low-level DsRed expression is punctuated with brighter, homozygous mutant nuclei (yellow arrowheads). (*F*) *y w Smg1^{32AP} FRT^{19A}/y w FRT^{19A}; e22c-GAL4, UAS-nlsDsRed2::SV40 3′UTR, UAS-FLP/+* L3 larva shows mosaic DsRed expression in epidermal nuclei: Low-level DsRed expression is punctuated with brighter, homozygous mutant nuclei (yellow arrowheads). Scale bars, (*A*,*B*) 1 mm; (*D*–*F*) 0.5 mm.

the SV40 3′ UTR can thus be used to monitor NMD pathway activity on a single-cell level, and this allowed us to perform a screen for genes involved in NMD using a genetic mosaic approach. Importantly, this mosaic approach allowed us to bypass the organismal lethality associated with strong loss of NMD function. Furthermore, by using a live-cell marker, we could score for NMD defects in F_1 generation animals, greatly increasing the screen throughput.

Our screen used the *e22c-GAL4* driver, which is expressed in the early embryonic ectoderm and later is restricted to ectodermally derived epithelium (Duffy et al. 1998), including the larval epidermis. This driver served two purposes in our screen (Fig. 1C). Expression of FLP recombinase driven by *e22c-GAL4* early in development produces animals with mosaicism in ectodermal derivatives. Later in development, *e22c-GAL4* drives expression of an NMD-sensitive fluorescent reporter in larval epidermal cells. Importantly, we found that this expression in epidermal cells is uniform (Fig. 1D), so differences in reporter levels caused by mutations in NMD genes could be readily identified.

To test our mosaic strategy, we crossed our mosaic generating reporter line to animals carrying known mutations in the core NMD components *Upf1* and *Upf2* (Fig. 1E; data not shown). Animals entirely mutant for null mutations in either of these genes show 100% early larval lethality and have strong defects in NMD (Metzstein and Krasnow 2006; Avery et al. 2011). We found that animals with tissue-directed mosaics of *Upf1* or *Upf2* were viable to the L3 larval stage, and we could easily detect increased expression in homozygous epidermal cells, confirming the feasibility of our approach for identifying organismal lethal mutations affecting NMD. We also found that mutants of the NMD auxiliary factor *Smg1*, which is viable and has weak defects in NMD (Chen et al. 2005; Metzstein and Krasnow 2006; Avery et al. 2011), produced an obvious mosaic phenotype (Fig. 1F). Thus, our strategy should allow us to isolate alleles of both essential and nonessential NMD genes.

We screened for mutations affecting NMD on the X chromosome and the right arm of the third chromosome using chemical mutagenesis and our mosaic approach (Fig. 2). While we identified our new mutants in mosaic animals, for characterization of their roles in the NMD pathway, we studied the alleles in nonmosaic, homozygous animals. In this way, we were able to determine the degree to which the new mutation is required for NMD. We were also able to assess whether the new mutation was lethal or viable in

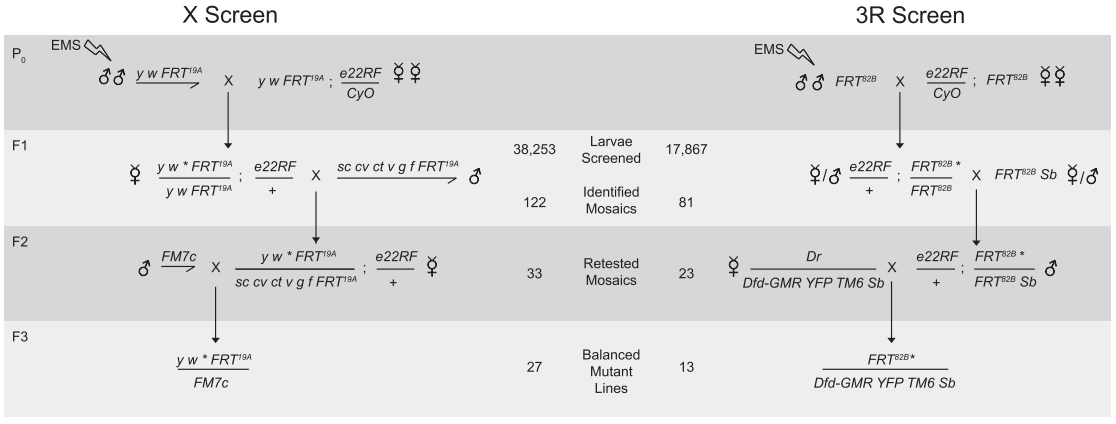


FIGURE 2. Screens to identify genes required for NMD on X and 3R. In the P_0, males carrying the appropriate FRT are mutagenized by treatment with EMS and mated *en masse* to reporter-containing females. In the F_1 generation, candidate NMD-mutant L3 larvae are identified by mosaic epidermal reporter expression, collected, and individually mated to flies containing marked FRTs. F_2 larvae are scored for transmission of the mosaic-expression phenotype, collected, and crossed to balancer-containing animals to establish stocks. (*Middle* column) Numbers of animals identified in each generation. (*e22RF*) Genotype *e22c-GAL4, UAS-nlsDsRed2::SV40 3′UTR, UAS-FLP*.

homozygous animals, giving insight into the requirement of the NMD pathway into organismal survival.

X-chromosome screen has isolated numerous alleles of known NMD genes

We identified 27 X-chromosome mutants, including both lethal and viable alleles (Table 1). Complementation tests revealed that all of these are alleles of the known NMD genes *Upf1*, *Upf2*, and *Smg1*, in which mutations had been previously isolated (Chen et al. 2005; Metzstein and Krasnow 2006). The most interesting phenotypic diversity we found was in our alleles of *Upf2*. When tested in homozygous animals, we found that six of these alleles were lethal and six were viable. This phenotypic difference provided us an opportunity to study the effects of *Upf2* mutations in vivo. Upf2 is thought to function primarily as a bridge between the other two core components of the NMD complex, Upf1 and Upf3. Binding of Upf2 to Upf3 has been mapped to a conserved MIF4G domain (Fig. 3A), while binding of Upf2 to Upf1 occurs primarily through a C-terminally localized domain, with a contribution from a domain close to the N terminus (He et al. 1997; Kadlec et al. 2004; Clerici et al. 2009). Upf2 also contains two other MIF4G domains of unknown function (Fig. 3A). To identify which domains of the Upf2 protein are necessary for viability, we sequenced the coding region of *Upf2* in eight of our alleles (four lethal and four viable) (Fig. 3A). Our analysis of these *Upf2* alleles revealed an interesting pattern: Lethal mutations are predicted to disrupt the second MIF4G domain but leave the other protein domains intact. Two previously characterized lethal alleles of *Upf2* are also predicted to disrupt the second MIF4G domain (Metzstein and Krasnow 2006); although these are also protein truncations, this disruption is not very specific. We did recover mutations that are predicted to disrupt the Upf2 interaction with Upf1 and that were not lethal. When we assayed NMD pathway activity in these viable mutants, we found that there was always residual pathway activity (Fig. 3B). These results suggest that pathway activity correlates with viability. This agrees with previous results showing that complete loss of the Upf1/Upf2 interaction is not compatible with viability (Avery et al. 2011). However, we cannot currently rule out that the second MIF4G domain has a cellular function independent of its role in NMD and disruption of this function also results in organismal

TABLE 1. 27 alleles recovered on the X chromosome

Upf1		*Upf2*		*Smg1*	
Allele	Viability	Allele	Viability	Allele	Viability
42	Lethal	*14-01-3M*	Lethal	*7-24-3*	Viable
A92A	Lethal	*3-01-3R*	Lethal	*8-9A*	Viable
8-7H	Lethal	*7-5A*	Lethal	*7-10-2*	Viable
14D	Lethal	*C41*	Lethal	*11-29-3-1*	Viable
		C51	Lethal	*C9*	Viable
		C2	Lethal	*C73*	Viable
		2-8A	Viable	*8D2*	Viable
		4-5F	Viable	*A6*	Viable
		C72	Viable	*A91*	Viable
		3D	Viable	*6B*	Viable
		C52	Viable	*11-2-10B*	Viable
		13D	Viable		

X-chromosome screen mutants. All alleles of *Upf1* are lethal, and all alleles of *Smg1* are viable, while alleles of *Upf2* can be either lethal or viable.

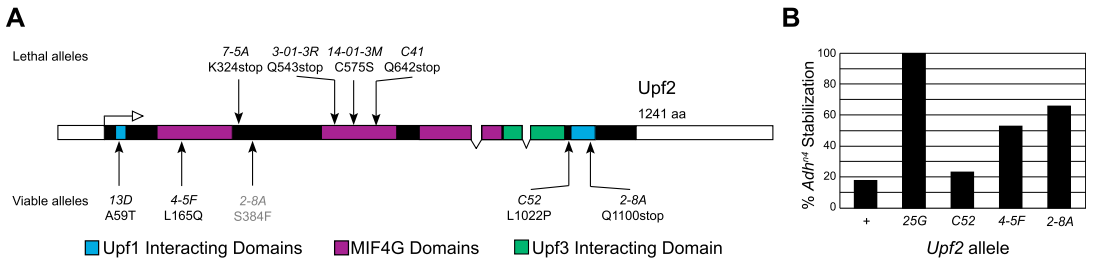


FIGURE 3. Characterization of *Upf2* alleles. (*A*) Mutations in *Upf2*. Alleles shown above the gene model are hemizygous lethal; those below the gene are hemizygous viable. *2-8A* contains two mutations: a late glutamine to stop (black), as well as a serine-to-phenylalanine missense mutation (gray). This latter change is in a nonconserved domain of Upf2, and we suspect it does not contribute to the *Upf2* defect. (Filled boxes) Coding sequences; (open boxes) 5′ and 3′ UTRs; (blue) bipartite Upf1 binding domain; (green) Upf3 binding domain; (purple) MIF4G domains (the third one of which overlaps the Upf3 binding domain); (open arrow) direction of transcription. (*B*) Effect of viable *Upf2* mutations on levels of the PTC-containing transcript Adh^{n4} (Q83Stop). Values are the proportion of Adh^{n4} transcript relative to Adh^{+} transcript normalized to $Upf2^{25G}$, in which Adh^{n4} is completely stabilized.

lethality. Finally, we noted that three of the four viable *Upf2* alleles were male sterile (data not shown), suggesting a role for Upf2 in the male germline (Table 2). Whether this role is directly related to its role in NMD has not yet been determined.

We obtained four alleles of *Upf1* but did not find that any of them are compatible with viability (Table 1), suggesting that even subtle perturbations in Upf1 structure are disruptive to protein function. In addition to *Upf1* and *Upf2*, we isolated 11 alleles of *Smg1*. It has previously been observed that complete loss of *Smg1* leads to no effect on viability and only a slight decrease in NMD function (Chen et al. 2005; Metzstein and Krasnow 2006). Indeed, we found that all 11 alleles were viable (Table 1).

Based on the large number of *Upf1*, *Upf2*, and *Smg1* alleles, we believe that we have saturated the X chromosome for genes required for NMD and that can be mutated by EMS, the chemical mutagen we used. Notably, we did not isolate any alleles of CG32533, the *Drosophila* homolog of the *C. elegans* gene *smgl-2* (Longman et al. 2007), which is located on the X chromosome. In addition to not identifying mutations in CG32533 in our screen, we also tested its function using an RNAi approach (A Chapin and MM Metzstein, unpubl.) and did not detect an NMD defect. Combined, these data suggest that *smgl-2* does not play a significant role in NMD in *Drosophila*. However, since EMS mutagenesis and RNAi knockdown do not affect all genes equally, it will be necessary to obtain mutations in *smgl-2* by some other means to definitively test the role of this gene in NMD in *Drosophila*.

Third chromosome screen has isolated first *Drosophila* alleles of *Smg6*

We screened the right arm of the third chromosome and identified 13 mutants. Similar to the X chromosome, after identifying the mutants with our mosaic approach, we studied the role of the genes affected in wholly mutant animals. By a complementation test with a chromosomal deficiency, we found that six of our 13 lines were likely to be mutations in *Smg6* (Supplemental Fig. S1). In all cases, the *Smg6* mutation was viable in *trans* to the deficiency. These six alleles also failed to complement each other (using the assay of reporter expression) (Supplemental Fig. S1). Three of the lines were viable as homozygotes. However, the remaining three lines were homozygous lethal, suggesting that extraneous lethal mutations are present on these chromosomes. The other seven lines complemented the deficiency for both viability and reporter enhancement and will be characterized further in a later study.

To identify the molecular changes in our candidate *Smg6* alleles, we sequenced the *Smg6* genomic locus in each of the six mutant lines. A single mutation affecting the *Smg6* gene was identified in each line (Fig. 4A; Table 3). The identified lesions are spread throughout the length of the *Smg6* gene and include premature termination codons, splice mutations, and a single missense change. This missense change,

TABLE 2. Molecular analysis of *Upf2* alleles

Allele	Viability	Mutation	Residue change	Male fertile
14-01-3M	Lethal	T1723A	C575S	N.A.
3-01-3R	Lethal	G1627A	Q543Stop	N.A.
7-5A	Lethal	A1694T	K324Stop	N.A.
C41	Lethal	G1924A	Q642Stop	N.A.
2-8A	Viable	C1876T	S384F	No
		C3298T	Q1100Stop	
4-5F	Viable	T494A	L165Q	No
C52	Viable	T3065C	L1022P	No
13D	Viable	G175A	A59T	Yes

For mutation positions, bases are numbered with the A of the start codon as 1, and introns are included. *7-5A* appears to be identical to the previously isolated *Upf2* allele *29AA* (Metzstein and Krasnow 2006).

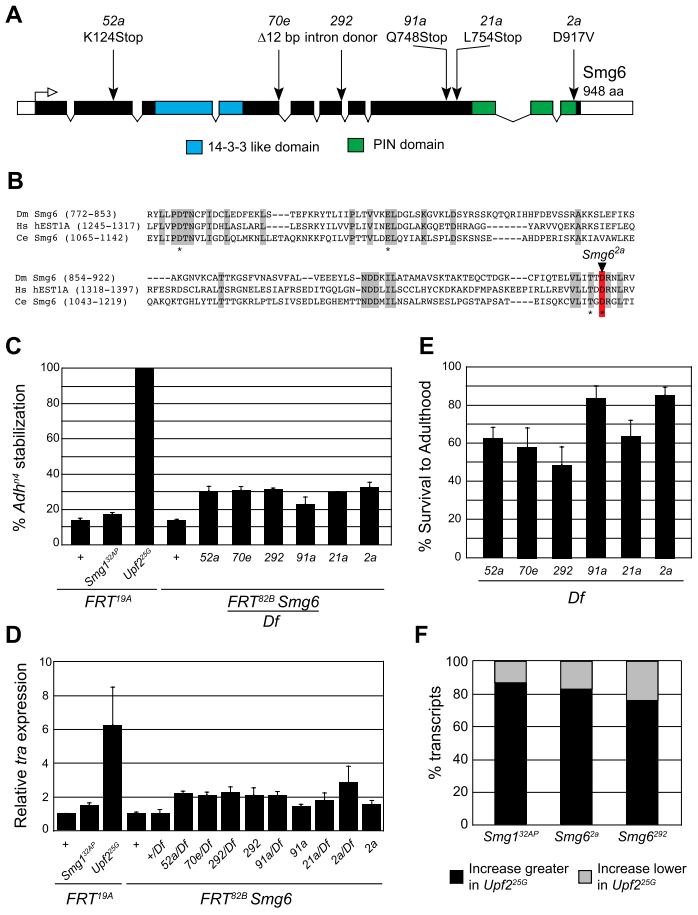


FIGURE 4. Molecular and phenotypic characterization of *Smg6* alleles. (*A*) Mutations in *Smg6*. (Filled boxes) Coding sequences; (open boxes) 5′ and 3′ UTRs; (blue) 14-3-3-like protein binding domain; (green) PIN endonucleolytic cleavage domain; (open arrow) direction of transcription. (*B*) Alignment of *Smg6* PIN domains. These highly conserved domains are representative of PIN domains found in Smg6 and other PIN-domain-containing proteins (Takeshita et al. 2007). (Dm) *D. melanogaster*; (Hs) *Homo sapiens*; (Ce) *C. elegans*. (*) Invariant catalytic residues; (gray shading) residues conserved in all three species; (red shading) aspartic acid residue altered to a valine in $Smg6^{2a}$. (*C*) Effect of *Smg6* mutations on levels of the PTC-containing transcript Adh^{n4} (Q83Stop). Values are the proportion of Adh^{n4} transcript relative to Adh^{+} transcript normalized to $Upf2^{25G}$, in which the Adh^{n4} is completely stabilized. Error bars indicate SD, $n = 2$. (*D*) Effect of *Smg6* mutations on levels of the endogenous NMD target *tra* in males. Levels for *Smg6* alleles are normalized to FRT^{82B}/Df animals; $Smg1^{32AP}/Y$ and $Upf2^{25G}/Y$ are normalized to FRT^{19A}/Y. Error bars indicate SEM, $n = 3$. (*E*) Percentage survival to adulthood of *Smg6* mutants. Each of the alleles are in *trans* to an *Smg6* deficiency. The percentage is calculated relative to balancer-containing siblings. Error bars: 95% confidence interval of the binomial distribution. (*F*) Microarray analysis of relative magnitude of NMD defects observed in $Upf2^{25G}$, $Smg1^{32AP}$, and two *Smg6* alleles, $Smg6^{2a}$ and $Smg6^{292}$. For each mutant genotype, we calculated transcript expression relative to its respective wild-type control. To compare the effect of different NMD genes with transcript regulation, we first extracted targets that increased more than twofold in all four of our NMD-mutant conditions, giving a list of 87 genes (Supplemental Table S1). We then compared the magnitude of the increase observed in $Smg1^{32AP}$, $Smg6^{2a}$, and $Smg6^{292}$ with that observed in $Upf2^{25G}$. (Black bars) Transcripts that increased more in $Upf2^{25G}$ relative to transcripts in *Smg1* or *Smg6*; (gray bars) transcripts that increase more in *Smg1* or *Smg6* alleles relative to $Upf2^{25G}$.

found in the allele $Smg6^{2a}$, is of particular interest, because it alters an invariantly conserved aspartate residue found in all PIN domains (Fig. 4B) and is known from structural studies to be part of the endonucleolytic catalytic core (Glavan et al. 2006; Takeshita et al. 2007).

Complete loss of *Smg6* still shows NMD-pathway activity in *Drosophila*

To characterize the role of *Drosophila Smg6* in NMD, we analyzed mRNA levels of the well-characterized NMD

TABLE 3. Molecular analysis of *Smg6* alleles

Allele	Base change	Residue change
52a	A535T	K124Stop
70e	GAgtgagtatat465Δ	N.A.
292	g558a	N.A.
91a	C2775T	Q748Stop
21a	T2794A	L754Stop
2a	A3580T	D917V

For mutation positions, bases are numbered with the A of the start codon as 1, and introns are included. *70e* is a 12-bp deletion that removes the last two bases of the fourth exon (uppercase) and 10 bp of the fourth intron (lowercase). *292* is a point mutation that converts the G of the sixth intron splice donor to an A.

target gene Adh^{n4}, which contains a nonsense mutation (Chia et al. 1987; Brogna 1999). Adh^{n4} has been shown to be a bona fide NMD target in both cell culture and whole animals (Brogna 1999; Gatfield et al. 2003; Chen et al. 2005; Metzstein and Krasnow 2006). Furthermore, *Adh* is expressed in most, if not all *Drosophila* tissues (Chintapalli et al. 2007), so by analyzing this gene we can assess the role played by NMD in diverse cell types.

Steady-state mRNA levels depend on mRNA stability and are thus a proxy for decay rates, and in Adh^{n4}/Adh^{+} animals, Adh^{n4} mRNA is expressed at ~10% of the amount of the Adh^{+} transcript, due to NMD-dependent destabilization of the mutant transcript (Brogna 1999; Chen et al. 2005; Metzstein and Krasnow 2006). We found that loss of *Smg6* function results in stabilization of the Adh^{n4} transcript to ~30% of the amount of Adh^{+} transcript (Fig. 4C). A similar degree of stabilization was observed in each of the six *Smg6* alleles in *trans* to a deficiency (Fig. 4C). In comparison, as previously reported, we found that *Smg1* mutants had only a very slight effect on Adh^{n4} levels (Chen et al. 2005; Metzstein and Krasnow 2006), whereas $Upf2^{25G}$, a strong hypomorphic allele of the core NMD component *Upf2* that essentially abolishes NMD, leads to complete stabilization of Adh^{n4} (Metzstein and Krasnow 2006). Loss of *Smg6* leads to considerably higher levels of Adh^{n4} than loss of *Smg1*, but significantly, stabilization of Adh^{n4} in *Smg6* is not nearly as strong as that observed in $Upf2^{25G}$. This suggests that *Smg6* is not absolutely required for NMD in vivo. We found similar results analyzing a PTC-containing allele of the developmental gene *eyes absent* (*eya*) (Supplemental Fig. S2A).

To determine if any of our *Smg6* alleles are definitively null, we used the classic genetic test of comparing homozygotes to the allele in *trans* to a deficiency. For this, we used quantitative reverse transcription coupled PCR (qRT-PCR) to measure the levels of the endogenous NMD-target gene *transformer* (*tra*) in *Smg6* mutant males (Fig. 4D) (this analysis was possible using the alleles *292*, *91a*, and *2a*, since these were not associated with any linked lethal mutation). Like *Adh*, *tra* is broadly expressed in *Drosophila* tissues and has been shown to be a bona fide NMD target in both intact animals (Metzstein and Krasnow 2006; Avery et al. 2011) and S2 cell culture (Rehwinkel et al. 2005).

Similar to what we observed for Adh^{n4}, we found that loss of *Smg6* function results in increased expression of *tra* to a level intermediate between *Smg1* and $Upf2^{25G}$. Of particular note, we observed that the degree of stabilization of *tra* was the same in $Smg6^{292}$ homozygotes and $Smg6^{292}$/*Df* hemizygotes, defining this as a null allele. Additionally, one of our other alleles, $Smg6^{52a}$, is predicted be a molecular null, because it contains an early stop codon truncating the coding region of *Smg6* before any known functional domains. Although we could not examine $Smg6^{52a}$ homozygotes due to a linked lethal mutation, we did find that $Smg6^{52a}$ displays a similar degree of *tra* stabilization in *trans* to a deficiency, as does $Smg6^{292}$. These results were confirmed using two other endogenous NMD targets, *Smg5* and *Oda* (Supplemental Fig. S2B; Gatfield and Izaurralde 2004; Rehwinkel et al. 2005; Avery et al. 2011). Thus, by both molecular and genetic criteria, we have isolated the first definitively null alleles of *Smg6*.

Interestingly, two of the *Smg6* alleles that specifically disrupt the PIN domain seem to retain residual activity. $Smg6^{91a}$, a nonsense mutation just preceding the PIN domain, and $Smg6^{2a}$, the missense mutation in the PIN domain, have lower levels of *tra* stabilization than null alleles, and this stabilization is increased when these alleles are placed in *trans* to a deficiency (Fig. 4D; Supplemental Fig. S2B). These results suggest that Smg6 has some activity independent of the PIN domain and thus presumably independent of its role as an endonuclease.

In general, the severity of NMD mutations and organismal viability are correlated in *Drosophila*. The NMD genes *Upf1* and *Upf2* are lethal as null alleles in *Drosophila*. The strong hypomorphic allele of $Upf2^{25G}$ is strongly disrupted in NMD function, and mutant animals are only partially viable, while a deletion of *Upf3* has moderate effects on NMD and is viable (Avery et al. 2011). Null alleles of *Smg1* have mild effects on NMD and are fully viable (Chen et al. 2005; Metzstein and Krasnow 2006). We therefore investigated the requirement for *Smg6* for viability by analyzing each of the *Smg6* alleles in *trans* to a deficiency for their survival to adulthood. In each case, loss of *Smg6* leads to only a moderate reduction in viability (Fig. 4E), further demonstrating that *Smg6* is not absolutely required for NMD in *Drosophila*.

To determine on a more global level whether loss of *Smg6* activity had less effect on the NMD pathway than loss of *Upf2*, we performed a microarray analysis. RNA was extracted from $Upf2^{25G}$, $Smg6^{2a}$, and $Smg6^{292}$ mutant animals and hybridized to microarrays representing the complete *Drosophila* transcriptome. We then tabulated genes that increased twofold over matched wild-type controls in all of our mutant conditions (Supplemental Table S1). We found that the large majority of

such genes had increased expression levels in $Upf2^{25G}$ mutants higher than the increase observed in *Smg6* mutants (Fig. 4F). From this, we conclude that, for the majority of endogenous NMD targets, *Smg6* mutants still possess residual NMD-pathway activity. When we tested $Smg1^{32AP}$ by the same assay, we found that it led to generally lower levels of transcript stabilization compared with *Smg6* (Fig. 4F). These results are consistent with our analysis of candidate transcripts, described above.

Another phenotype associated with loss of NMD function in *Drosophila* is impairment of cell proliferation, both in S2 cell culture as well as in vivo (Rehwinkel et al. 2005; Metzstein and Krasnow 2006; Avery et al. 2011). We examined cell proliferation mosaic analysis in the eye, a technique that permits comparison of homozygous mutant cell growth in the presence or absence of wild-type cells (Stowers and Schwarz 1999). We find that, similar to other NMD mutants (Metzstein and Krasnow 2006; Avery et al. 2011), *Smg6* mutant cells are defective in cell proliferation when in competition with wild-type cells, as assayed by examination of mitotic clones within the eye (Supplemental Fig. S4A–D). However, when competing cells are removed, *Smg6* mutant cells can essentially proliferate as well as wild-type cells (Supplemental Fig. S4E–H). These data show that while *Smg6* is dispensable for viability, it is still required for aspects of cell function in vivo.

Conclusions

We have isolated point mutations in *Drosophila Smg6*, which are the first mutant alleles of *Smg6* characterized in any organism since the initial identification of this gene in *C. elegans* (Hodgkin et al. 1989). Since a number of our *Smg6* alleles clearly lack all function, this has allowed us to assess the contribution of this gene to the NMD process in vivo. We find that in *Drosophila*, Smg6 has only a moderate role in targeting PTC-containing transcripts for degradation. In comparison, mutations in core NMD components, such as *Upf2*, lead to complete stabilization of NMD targets in vivo (Metzstein and Krasnow 2006; Avery et al. 2011). Our results contrast with previous analysis of *Smg6* in *Drosophila* S2 cells using RNAi-mediated knockdown, in which inactivation of *Smg6* appeared to result in as strong an effect on NMD targets as inactivation of other components of the NMD pathway (Gatfield et al. 2003; Rehwinkel et al. 2005). One explanation for this contradiction is that it is not possible to assess the exact degree of gene inactivation in RNAi knockdown experiments, making comparisons of the requirements of different genes in a process difficult to assess.

It is interesting to compare our results of *Smg6* mutations in vivo with the study of this gene in other species. In *C. elegans*, the organism in which *Smg6* was first identified, the NMD process seems to be as dependent on *Smg6* as any of the other NMD genes (Hodgkin et al. 1989). However, because no description of the molecular nature of *C. elegans Smg6* alleles has been published, it is unknown if any of the known alleles are null. Indeed, there is some evidence that null alleles of *C. elegans smg-6* are not viable, thus the alleles allowing viable NMD-defective animals may indeed possess residual function (Cali and Anderson 1998). In zebrafish, *Smg6* has been analyzed by morpholino knockdown in intact fish and by expression of an *Smg6* dominant-negative construct in fish cell lines (Wittkopp et al. 2009). The morpholino knockdown behaves very similarly to knockdown of other NMD components, leading to early lethality with distinct patterning and neural defects, suggesting that *Smg6* has similar roles to these other components in vivo. However, because direct measurement of NMD-pathway activity was not performed in the intact *Smg6* morphants, it is possible that there is still residual pathway activity, as observed in *Drosophila*. Measurements of NMD-pathway activity were analyzed in fish cell culture and clearly show a defect in NMD. Whether this defect is equivalent to complete NMD-pathway loss was not analyzed.

Smg6 is an endonuclease that cleaves target mRNAs around the site of the premature termination codon (Huntzinger et al. 2008), and our data support this as an important Smg6 function. All of our alleles of *Smg6* are predicted to disrupt the C-terminal endonucleolytic PIN domain, including one that is a point mutation affecting a key catalytic residue. However, since *Smg6*-null mutants still show considerable NMD-pathway activity, this suggests that in intact *Drosophila* there exists an Smg6-independent degradation pathway. In yeast, NMD targets are thought to be degraded by decapping followed by 5′–3′ exonucleolytic degradation (Muhlrad and Parker 1994; He et al. 2003). Inactivation of this pathway was shown not to have an effect on NMD-mediated degradation in S2 cells (Gatfield et al. 2003), but one possibility is that this mechanism is only used in the absence of Smg6-mediated cleavage. Indeed, mammalian cells are known to use multiple decay pathways (Chen and Shyu 2003; Lejeune et al. 2003; Huang et al. 2011). *Drosophila* genetics may provide a powerful system for identifying such alternative degradation pathways, because mutations in such genes would be predicted to act as enhancers of the *Smg6* phenotype.

Like Smg6, Smg1 is not absolutely required for NMD in *Drosophila* (Chen et al. 2005; Metzstein and Krasnow 2006). Smg1 is a PIKK-family kinase that phosphorylates the core NMD component Upf1 to provide a binding platform for Smg6 (Kashima et al. 2006; Okada-Katsuhata et al. 2012). If such binding were absolutely necessary for Smg6 activity, we would expect that *Smg1* mutations would have at least as strong an NMD defect as *Smg6* mutations. However, we find just the opposite: Loss of Smg6 activity leads to a significantly stronger NMD defect than loss of Smg1, suggesting that Smg6 has activity independent of

Smg1-mediated phosphorylation of Upf1. In mammalian cells, Smg6 can interact with the NMD complex independent of binding to phosphorylated Upf1 through interactions between two N-terminal exon junction complex binding domains (Kashima et al. 2010). *Drosophila* is thought to lack these domains, but it remains possible that Upf1-phosphorylation-independent interactions can still occur. Alternatively, the *Drosophila* genome may contain an alternative kinase redundant with Smg1 that can phosphorylate Upf1. This could explain why Smg1 has a relatively weak affect on NMD in intact *Drosophila* and provide an alternative mechanism of Smg6 entry into the NMD complex. It should be possible to identify such a kinase by testing other kinases in the *Drosophila* genome for synergistic effects with *Smg1* mutants on the NMD pathway.

Analysis of NMD in mammalian cells has suggested that NMD is not a linear pathway, but potentially different branches act in parallel to robustly regulate the abundance of target mRNAs. For instance, there exist Upf3-dependent and Upf3-independent NMD pathways targeting certain nonsense transcripts in mammalian cells (Chan et al. 2007; Yepiskoposyan et al. 2011). Our genetic screen and molecular analysis presented here has also revealed that multiple mechanisms function to regulate NMD in *Drosophila*. Finally, the efficient screening method for NMD genes presented here should be applicable to the entire *Drosophila* genome, providing the opportunity to identify all of the components contributing to these parallel mechanisms.

MATERIALS AND METHODS

Fly stocks

Drosophila melanogaster lines were raised at 25°C on cornmeal/dextrose medium using standard conditions. Control chromosomes used for the screens were *y w* FRT^{19A} and FRT^{82B} (Xu and Rubin 1993); the *e22c-GAL4* and *UAS-FLP* stocks (Duffy et al. 1998) were used in all screen stocks to induce epidermal mosaics using the GAL4/UAS system (Brand and Perrimon 1993) and FLP/FRT systems (Xu and Rubin 1993). The *nlsDsRed2::SV40 3′UTR* reporter transgene was provided by M. Galko (M.D. Anderson Cancer Center). *e22c-GAL4, UAS-nlsDsRed2::SV40 3′UTR, UAS-FLP* was built using standard recombination methods (MM Metzstein, unpubl.). NMD alleles $Upf2^{25G}$, $Smg1^{32AP}$, and $Upf1^{26A}$ have been described previously (Metzstein and Krasnow 2006). The *breathless* promoter *btl-Gal4* was used in combination with *UAS-GFP* as the reporter in the X-chromosome complementation tests (Shiga et al. 1996). For the balancer chromosomes, we used *w*;* ry^{506} Dr^1*/TM6B, P{Dfd-EYFP}* Sb^1 ca^1 (Le et al. 2006) and *FM7i, P{ActGFP}JMR3* (Reichhart and Ferrandon 1998). Complementation tests and deficiency experiments for *Smg6* were performed using *Df(3R)ED6220/TM6c Tb, Sb* (Ryder et al. 2007). Eye clones were generated using *w; ey-FLP;* FRT^{82B} *P{w+ ry+}90E* and *y w; ey-GAL4, UAS-FLP;* FRT^{82B} *GMR-hid/TM6* (Stowers and Schwarz 1999). For NMD tests, we used the PTC-containing alleles Adh^{nt4} (Chia et al. 1987) and $eya^{cli\text{-}IID}$ (Bui et al. 2000).

Mutagenesis and screen scheme

For mutagenesis, isogenized males were aged 1–3 d, starved overnight, and fed 25 mM ethane methylsulfonate in 1% sucrose for 24 h (Newsome et al. 2000). For the X-chromosome screen, *y w* FRT^{19A}*/Y* mutagenized males were mated *en masse* to *y w* FRT^{19A}*; e22c-GAL4, UAS-nlsDsRed2::SV40 3′UTR, UAS-FLP/CyO* virgins. F_1 wandering third-instar larvae were collected in glycerol and analyzed using a Leica MZ16 fluorescence stereomicroscope. NMD mutants were identified by mosaic enhancement of the DsRed reporter in the larval epidermis. Candidate mutant F_1 larvae were collected, rinsed in 1× PBS, placed in individual vials, and allowed to complete development. Surviving F_1's were mated to *sc cv ct v g f* FRT^{19A}*/Y*, and F_2 progeny were retested for mosaic enhancement of DsRed expression; mosaic animals were collected and mated to *FM7c/Y* males to establish balanced mutant lines. We screened about 38,000 DsRed-expressing F_1 animals. Of these, we expected 50% to be females, which could exhibit mosaicism (the others are hemizygous males and thus would not be expected to undergo FLP-mediated recombination). We identified 122 mosaic larvae from which we recovered 27 balanced mutant lines, representing a recovery rate of ∼22%.

For the third chromosome, mutagenized *w/Y;* FRT^{82B} males were mated *en masse* to *e22c-GAL4, UAS-nlsDsRed2::SV40 3′UTR, UAS-FLP /CyO;* FRT^{82B} virgins, and F_1 animals were scored and collected as described above. Candidate mutants were retested by mating to *y w;* FRT^{82B} Sb^1*/ TM6B ry* Tb^+ animals of the opposite sex, and male progeny that retested were collected and mated to *w; Kr/CyO; D/ TM6c Tb, Sb* or to *w*;* ry^{506} Dr^1*/TM6B, P{Dfd-EYFP},* Sb^1 ca^1 to establish mutant lines. We scored about 17,000 F_1's (both males and females would be expected to be mosaics), and we identified 81 mosaic larvae. Of these we were able to establish 13 balanced mutant lines, representing a recovery rate of ∼16%.

Complementation tests

X

There are three known NMD genes on the X chromosome with previously isolated alleles (Metzstein and Krasnow 2006). Null alleles of *Upf1* and *Upf2* are lethal, whereas null alleles of *Smg1* are viable (Chen et al. 2005; Metzstein and Krasnow 2006). Complementation tests were performed using these alleles. For *Upf1* complementation tests, *y w* FRT^{19A} **/FM7c* (where * indicates EMS-induced mutation) virgins were mated to *y w* $Upf1^{13D}$ FRT^{19A}*/Dp(1;Y)BSC1[*$Upf1^+$*]* males, and F_1 progeny were scored for the absence of B^+ females, which indicates failure to complement *Upf1* for lethality. In cases in which complementation for lethality occurred, the mutant lines were then crossed to *y w* $Upf1^{13D}$ FRT^{19A}*/Dp(1;Y)BSC1[*$Upf1^+$*]; Btl-Gal4, UAS-GFP::SV40 3′UTR/+* males, and F_1 female third-instar larvae were scored for enhancement of GFP. Failure to complement this phenotype would have indicated that the candidate gene is a viable *Upf1* hypomorph. However, we did not find any alleles of this class.

For *Upf2* complementation tests, *y w FRT*/FM7c* virgins were crossed to *y w* $Upf2^{14J}$ *v g f* FRT^{19A}*/Y; P{*w^+*,* $Upf2^+$*}/+* males (Metzstein and Krasnow 2006), and F_1 animals were scored for the absence of *w* B^+ females, which indicates failure to complement *Upf2* for lethality. In cases in which complementation occurred,

the mutant lines were crossed to *y w Upf2^{14J} v g f FRT^{19A}/Y; P{w$^+$, Upf2$^+$}/Btl-Gal4, UAS-GFP::SV40 3′UTR* males, and F_1 female third-instar larvae were scored for GFP enhancement. Failure to complement indicates that the *Upf2* allele is a viable hypomorph. For *Smg1* complementation tests, females from mutant lines were crossed to *y w Smg1^{32AP} FRT^{19A}/Y; Btl-Gal4, UAS-GFP::SV40 3′UTR/+* male larvae and were scored for GFP enhancement, indicating failure to complement *Smg1*.

3R

Complementation tests were performed by crossing *e22c-GAL4, UAS-nlsDsRed2::SV40 3′UTR, UAS-FLP/CyO; FRT^{82B}*/TM6c Tb, Sb* females to *Df(3R)ED6220/TM6c Tb, Sb* animals, and third-instar larvae were scored for DsRed enhancement. Lines that failed to complement *Smg6* were intercrossed to determine if the *Smg6* mutations were lethal or had additional linked lethals.

RNA isolation and quantitative RT-PCR

Total RNA was isolated from five to 10 flash-frozen samples of adult males using TRIzol reagent (Invitrogen) and phase-lock tubes (5-prime), and the RNeasy mini kit (QIAGEN). Genomic contamination was reduced by on-column DNase treatment (Ambion). RNA concentration was determined using a spectrophotometer and normalized for reverse transcription using the MuMulv reverse transcriptase Retroscript Kit (Ambion). Quantitative RT-PCR was performed on a Bio-Rad iCycler thermocycler using SYBR Green Supermix (Bio-Rad). All experimental reactions were performed in triplicate with a minimum of three biological replicates. For quantitation, samples were normalized to *RpL32*.

Analysis of *Adh^{n4}*

Adh^{n4}/CyO animals were mated to *Df(3R)ED6220/TM6B, P{Dfd-EYFP} Sb^1 ca^1*. Males in the F_1 generation *Adh^{n4}/Adh$^+$; Df(3R)ED6220/+* were collected and mated to virgin *FRT^{82B}/TM6B, P{Dfd-EYFP} Sb^1 ca^1* or *FRT^{82B} Smg6*/ TM6B, P{Dfd-EYFP} Sb^1 ca^1* flies. Animals of genotype *Adh^{n4}/Adh$^+$; FRT^{82B} Smg6*/Df(3R)ED6220* and *Adh^{n4}/Adh$^+$; FRT^{82B}/ Df(3R)ED6220* were collected by genotyping for the presence of the *Adh^{n4}*. Animals of genotypes *Smg1^{32AP}/Y; Adh^{n4}/Adh$^+$* and *Upf2^{25G}/Y; Adh^{n4}/Adh$^+$* were collected in parallel. Two biological replicates were collected for each genotype, and RNA isolation of five to eight adult males was performed as described above. Samples were normalized for RNA concentration and reverse-transcribed. The cDNA from each sample was used as template DNA in a PCR amplification of the Adh transcript. To distinguish the *Adh$^+$* and *Adh^{n4}* alleles, the PCR product was digested using PvuII and run on a 1% agarose gel (Supplemental Fig. S3). The gels were imaged using an Alpha Innotech Gel Doc System and quantitated using ImageJ (Abràmoff et al. 2004).

Analysis of *eya$^{cli\text{-}IID}$*

eya$^{cli\text{-}IID}$ cn^1 bw^1 sp^1/CyO males were mated to *e22c-GAL4, UAS-nlsDsRed2::SV40 3′UTR/CyO; FRT^{82B}**, where * represents *Smg6$^+$*, *Smg6^{2a}*, or *Smg6^{292}* virgins. Male progeny of genotype *eya$^{cli\text{-}IID}$ cn^1 bw^1 sp^1/ e22c-GAL4, UAS-nlsDsRed2::SV40 3′UTR; FRT^{82B}*/+* were mated to virgins of genotype *Btl-Gal4, UAS-GFP::SV40 3′UTR/+; Df(3R)ED6220/+*. In the next generation, we collected larvae of genotype *eya$^{cli\text{-}IID}$ cn^1 bw^1 sp^1/ Btl-Gal4, UAS-GFP::SV40 3′UTR; FRT^{82B} */Df(3R)ED6220* based on increased GFP expression (due to the *Smg6* mutation) and a lack of DsRed expression. A similar strategy was used on the X chromosome to collect larvae that were either *y w FRT^{19A}* or *y w Upf2^{25G} FRT^{19A} /Y; eya$^{cli\text{-}IID}$ cn^1 bw^1 sp^1/ e22c-GAL4, UAS-nlsDsRed2::SV40 3′UTR*. RNA was isolated as described above, reverse-transcribed, and amplified using primers surrounding the *eya$^{cli\text{-}IID}$* PTC mutation. We determined the sequence of the RT-PCR product. *eya$^{cli\text{-}IID}$* stabilization was determined by measuring heights of the *eya$^+$* and *eya$^{cli\text{-}IID}$* sequencing peaks.

Microarray analysis

RNA was isolated from male larvae collected from 0 to 4 h after the molt to third instar. To identify genotypes at this stage, we used balancers marked with fluorescent reporters: *ActGFP* for the X (Reichhart and Ferrandon 1998) and *Dfd-EYFP* for the third (Le et al. 2006). For *Smg6* experiments, we crossed *FM7i, P{ActGFP}JMR3/Y; Df(3R)ED6220/ TM6B, P{Dfd-EYFP} Sb^1 ca^1* males to either *w; FRT^{82B}/ TM6B, P{Dfd-EYFP} Sb^1 ca^1* or *w; Smg6^{2a}/ TM6B, P{Dfd-EYFP} Sb^1 ca^1*, or *w; Smg6^{292}/ TM6B, P{Dfd-EYFP} Sb^1 ca^1* virgins. For the X-chromosome NMD genes, we crossed *FM7i, P{ActGFP}JMR3/Y* males to either *y w FRT^{19A}* or *y w Smg1^{32AP} FRT^{19A}*, or *y w Upf2^{25G} FRT^{19A}/FM7i, P{ActGFP}JMR3* virgins. Progeny from these crosses were staged, and the appropriate genotypes were collected by scoring for the absence of the fluorescent balancers. RNA from two biological replicates of each genotype was reverse-transcribed, and the cDNA was labeled with either Cy3 or Cy5 and hybridized to Agilent 44K 2-color arrays. Data files were produced using Agilent's feature extraction software version 10.5. We normalized each array to total signal.

Smg6 viability analysis

To test the viability of *Smg6* alleles, virgin *FRT^{82B}* or *FRT^{82B} Smg6*/ TM6B, P{Dfd-EYFP} Sb^1 ca^1* flies were mated to *Df(3R)ED6220/ TM6B, P{Dfd-EYFP} Sb^1 ca^1* males, and progeny adults were analyzed for the presence and absence of the *Sb* marker associated with the balancer chromosome.

Eye clones

Mosaic eyes were generated by crossing *w; ey-FLP; FRT^{82B} P{w+ ry+}90E* virgins to *FRT^{82B}* or *FRT^{82B}** (where * is EMS-generated mutation) males. Images at multiple planes were taken on an Axioskop2 plus (Zeiss) microscope and analyzed using an ImageJ stack analyzer and a stack focuser plug-in. Entirely mutant eyes were generated by the *GMR-hid* technique (Stowers and Schwarz 1999).

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank Michael Galko for *UAS-nls:DsRed* transgenes; Alex Chapin and Jon Nelson for unpublished data; Shanhong Cheng,

Sarah Elliott, and Deborah Johnson for assistance in screening; Barry Moore and Mark Yandell for help with array bioinformatics; the Metzstein and Thummel laboratories for useful discussions; and Gillian Stanfield, Carl Thummel, and members of the Metzstein laboratory for their comments on the manuscript. Fly stocks were obtained from the Bloomington *Drosophila* Stock Center. This work was supported by National Institutes of Health (NIH) University of Utah Developmental Biology Training Grant 5T32-HD07491 (to K.A.F.) and NIH 1R01GM084011 and a March of Dimes Award 5-FY07-664 (to M.M.M).

Received February 20, 2012; accepted June 7, 2012.

REFERENCES

Abràmoff MD, Magalhães PJ, Ram SJ. 2004. Image processing with ImageJ. *Biophotonics international* **11:** 36–42.

Amrani N, Ganesan R, Kervestin S, Mangus DA, Ghosh S, Jacobson A. 2004. A *faux* 3′-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature* **432:** 112–118.

Anastasaki C, Longman D, Capper A, Patton EE, Cáceres JF. 2011. Dhx34 and Nbas function in the NMD pathway and are required for embryonic development in zebrafish. *Nucleic Acids Res* **39:** 3686–3694.

Aronoff R, Baran R, Hodgkin J. 2001. Molecular identification of *smg-4*, required for mRNA surveillance in *C. elegans*. *Gene* **268:** 153–164.

Avery P, Vicente-Crespo M, Francis D, Nashchekina O, Alonso CR, Palacios IM. 2011. *Drosophila* Upf1 and Upf2 loss of function inhibits cell growth and causes animal death in a Upf3-independent manner. *RNA* **17:** 624–638.

Brand AH, Perrimon N. 1993. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118:** 401–415.

Brogna S. 1999. Nonsense mutations in the alcohol dehydrogenase gene of *Drosophila melanogaster* correlate with an abnormal 3′ end processing of the corresponding pre-mRNA. *RNA* **5:** 562–573.

Bui QT, Zimmerman JE, Liu H, Bonini NM. 2000. Molecular analysis of Drosophila *eyes absent* mutants reveals features of the conserved Eya domain. *Genetics* **155:** 709–720.

Cali BM, Anderson P. 1998. mRNA surveillance mitigates genetic dominance in *Caenorhabditis elegans*. *Mol Gen Genet* **260:** 176–184.

Cali BM, Kuchma SL, Latham J, Anderson P. 1999. *smg-7* is required for mRNA surveillance in *Caenorhabditis elegans*. *Genetics* **151:** 605–616.

Chan W-K, Huang L, Gudikote JP, Chang Y-F, Imam JS, MacLean JA, Wilkinson MF. 2007. An alternative branch of the nonsense-mediated decay pathway. *EMBO J* **26:** 1820–1830.

Chang Y-F, Imam JS, Wilkinson MF. 2007. The nonsense-mediated decay RNA surveillance pathway. *Annu Rev Biochem* **76:** 51–74.

Chen C-YA, Shyu A-B. 2003. Rapid deadenylation triggered by a nonsense codon precedes decay of the RNA body in a mammalian cytoplasmic nonsense-mediated decay pathway. *Mol Cell Biol* **23:** 4805–4813.

Chen Z, Smith KR, Batterham P, Robin C. 2005. *Smg1* nonsense mutations do not abolish nonsense-mediated mRNA decay in *Drosophila melanogaster*. *Genetics* **171:** 403–406.

Chia W, Savakis C, Karp R, Ashburner M. 1987. Adh^{n4} of *Drosophila melanogaster* is a nonsense mutation. *Nucleic Acids Res* **15:** 3931. doi: 10.1093/nar/15.9.3931.

Chintapalli VR, Wang J, Dow JAT. 2007. Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nat Genet* **39:** 715–720.

Chiu S-Y, Serin G, Ohara O, Maquat LE. 2003. Characterization of human Smg5/7a: A protein with similarities to *Caenorhabditis elegans* SMG5 and SMG7 that functions in the dephosphorylation of Upf1. *RNA* **9:** 77–87.

Clerici M, Mourão A, Gutsche I, Gehring NH, Hentze MW, Kulozik A, Kadlec J, Sattler M, Cusack S. 2009. Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2. *EMBO J* **28:** 2293–2306.

Cui Y, Hagan KW, Zhang S, Peltz SW. 1995. Identification and characterization of genes that are required for the accelerated degradation of mRNAs containing a premature translational termination codon. *Genes Dev* **9:** 423–436.

Duffy JB, Harrison DA, Perrimon N. 1998. Identifying loci required for follicular patterning using directed mosaics. *Development* **125:** 2263–2271.

Eberle AB, Stalder L, Mathys H, Zamudio Orozco R, Mühlemann O. 2008. Posttranscriptional gene regulation by spatial rearrangement of the 3′ untranslated region. *PLoS Biol* **6:** e92. doi: 10.1371/journal.pbio.0060092.

Eberle AB, Lykke-Andersen S, Mühlemann O, Jensen TH. 2009. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nat Struct Mol Biol* **16:** 49–55.

Fukuhara N, Ebert J, Unterholzner L, Lindner D, Izaurralde E, Conti E. 2005. SMG7 is a 14-3-3-like adaptor in the nonsense-mediated mRNA decay pathway. *Mol Cell* **17:** 537–547.

Gatfield D, Izaurralde E. 2004. Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in *Drosophila*. *Nature* **429:** 575–578.

Gatfield D, Unterholzner L, Ciccarelli FD, Bork P, Izaurralde E. 2003. Nonsense-mediated mRNA decay in *Drosophila*: At the intersection of the yeast and mammalian pathways. *EMBO J* **22:** 3960–3970.

Glavan F, Behm-Ansmant I, Izaurralde E, Conti E. 2006. Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. *EMBO J* **25:** 5117–5125.

Grimson A, O'Connor S, Newman CL, Anderson P. 2004. SMG-1 is a phosphatidylinositol kinase-related protein kinase required for nonsense-mediated mRNA decay in *Caenorhabditis elegans*. *Mol Cell Biol* **24:** 7483–7490.

He F, Jacobson A. 1995. Identification of a novel component of the nonsense-mediated mRNA decay pathway by use of an interacting protein screen. *Genes Dev* **9:** 437–454.

He F, Brown AH, Jacobson A. 1997. Upf1p, Nmd2p, and Upf3p are interacting components of the yeast nonsense-mediated mRNA decay pathway. *Mol Cell Biol* **17:** 1580–1594.

He F, Li X, Spatrick P, Casillo R, Dong S, Jacobson A. 2003. Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5′ to 3′ mRNA decay pathways in yeast. *Mol Cell* **12:** 1439–1452.

Hodgkin J, Papp A, Pulak R, Ambros V, Anderson P. 1989. A new kind of informational suppression in the nematode *Caenorhabditis elegans*. *Genetics* **123:** 301–313.

Huang L, Lou C-H, Chan W, Shum EY, Shao A, Stone E, Karam R, Song H-W, Wilkinson MF. 2011. RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD. *Mol Cell* **43:** 950–961.

Huntzinger E, Kashima I, Fauser M, Saulière J, Izaurralde E. 2008. SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA* **14:** 2609–2617.

Hwang J, Maquat LE. 2011. Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: To die or not to die, that is the question. *Curr Opin Genet Dev* **21:** 422–430.

Kadlec J, Izaurralde E, Cusack S. 2004. The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3. *Nat Struct Mol Biol* **11:** 330–337.

Kashima I, Yamashita A, Izumi N, Kataoka N, Morishita R, Hoshino S, Ohno M, Dreyfuss G, Ohno S. 2006. Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes Dev* **20:** 355–367.

Kashima I, Jonas S, Jayachandran U, Buchwald G, Conti E, Lupas AN, Izaurralde E. 2010. SMG6 interacts with the exon junction complex via two conserved EJC-binding motifs (EBMs) required for nonsense-mediated mRNA decay. *Genes Dev* **24:** 2440–2450.

Kerényi Z, Mérai Z, Hiripi L, Benkovics A, Gyula P, Lacomme C, Barta E, Nagy F, Silhavy D. 2008. Inter-kingdom conservation of mechanism of nonsense-mediated mRNA decay. *EMBO J* **27:** 1585–1595.

Kertész S, Kerényi Z, Mérai Z, Bartos I, Pálfy T, Barta E, Silhavy D. 2006. Both introns and long 3′-UTRs operate as *cis*-acting elements to trigger nonsense-mediated decay in plants. *Nucleic Acids Res* **34:** 6147–6157.

Le T, Liang Z, Patel H, Yu MH, Sivasubramaniam G, Slovitt M, Tanentzapf G, Mohanty N, Paul SM, Wu VM, et al. 2006. A new family of *Drosophila* balancer chromosomes with a w^{-}*dfd*-GMR yellow fluorescent protein marker. *Genetics* **174:** 2255–2257.

Lee BS, Culbertson MR. 1995. Identification of an additional gene required for eukaryotic nonsense mRNA turnover. *Proc Natl Acad Sci* **92:** 10354–10358.

Leeds P, Peltz SW, Jacobson A, Culbertson MR. 1991. The product of the yeast *UPF1* gene is required for rapid turnover of mRNAs containing a premature translational termination codon. *Genes Dev* **5:** 2303–2314.

Leeds P, Wood JM, Lee BS, Culbertson MR. 1992. Gene products that promote mRNA turnover in *Saccharomyces cerevisiae*. *Mol Cell Biol* **12:** 2165–2177.

Lejeune F, Maquat LE. 2005. Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. *Curr Opin Cell Biol* **17:** 309–315.

Lejeune F, Li X, Maquat LE. 2003. Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. *Mol Cell* **12:** 675–687.

Longman D, Plasterk RHA, Johnstone IL, Cáceres JF. 2007. Mechanistic insights and identification of two novel factors in the *C. elegans* NMD pathway. *Genes Dev* **21:** 1075–1085.

Luke B, Azzalin CM, Hug N, Deplazes A, Peter M, Lingner J. 2007. *Saccharomyces cerevisiae* Ebs1p is a putative ortholog of human Smg7 and promotes nonsense-mediated mRNA decay. *Nucleic Acids Res* **35:** 7688–7697.

Metzstein MM, Krasnow MA. 2006. Functions of the nonsense-mediated mRNA decay pathway in *Drosophila* development. *PLoS Genet* **2:** e180. doi: 10.1371/journal.pgen.0020180.

Muhlrad D, Parker R. 1994. Premature translational termination triggers mRNA decapping. *Nature* **370:** 578–581.

Newsome TP, Asling B, Dickson BJ. 2000. Analysis of *Drosophila* photoreceptor axon guidance in eye-specific mosaics. *Development* **127:** 851–860.

Okada-Katsuhata Y, Yamashita A, Kutsuzawa K, Izumi N, Hirahara F, Ohno S. 2012. N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. *Nucleic Acids Res* **40:** 1251–1266.

Page MF, Carr B, Anders KR, Grimson A, Anderson P. 1999. SMG-2 is a phosphorylated protein required for mRNA surveillance in *Caenorhabditis elegans* and related to Upf1p of yeast. *Mol Cell Biol* **19:** 5943–5951.

Rehwinkel J, Letunic I, Raes J, Bork P, Izaurralde E. 2005. Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. *RNA* **11:** 1530–1544.

Reichhart JM, Ferrandon D. 1998. Green balancers. *Drosoph Inf Serv* **81:** 201–202.

Ryder E, Ashburner M, Bautista-Llacer R, Drummond J, Webster J, Johnson G, Morley T, Chan YS, Blows F, Coulson D, et al. 2007. The DrosDel deletion collection: A *Drosophila* genomewide chromosomal deficiency resource. *Genetics* **177:** 615–629.

Serin G, Gersappe A, Black JD, Aronoff R, Maquat LE. 2001. Identification and characterization of human orthologues to *Saccharomyces cerevisiae* Upf2 protein and Upf3 protein (*Caenorhabditis elegans* SMG-4). *Mol Cell Biol* **21:** 209–223.

Shiga Y, Tanaka-Matakatsu M, Hayashi S. 1996. A nuclear GFP/β-galactosidase fusion protein as a marker for morphogenesis in living *Drosophila*. *Dev Growth Differ* **38:** 99–106.

Stowers RS, Schwarz TL. 1999. A genetic method for generating *Drosophila* eyes composed exclusively of mitotic clones of a single genotype. *Genetics* **152:** 1631–1639.

Swisher KD, Parker R. 2011. Interactions between Upf1 and the decapping factors Edc3 and Pat1 in *Saccharomyces cerevisiae*. *PLoS ONE* **6:** e26547. doi: 10.1371/journal.pone.0026547.

Takeshita D, Zenno S, Lee WC, Saigo K, Tanokura M. 2007. Crystal structure of the PIN domain of human telomerase-associated protein EST1A. *Proteins* **68:** 980–989.

Wittkopp N, Huntzinger E, Weiler C, Sauliere J, Schmidt S, Sonawane M, Izaurralde E. 2009. Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic development and survival. *Mol Cell Biol* **29:** 3517–3528.

Xu T, Rubin GM. 1993. Analysis of genetic mosaics in developing and adult *Drosophila* tissues. *Development* **117:** 1223–1237.

Yepiskoposyan H, Aeschimann F, Nilsson D, Okoniewski M, Mühlemann O. 2011. Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. *RNA* **17:** 2108–2118.

CHAPTER 3

MAPPING AND PRELIMINARY ANALYSIS OF NEW NMD MUTANTS

Introduction

Biochemical analysis of nonsense mediated decay (NMD) has relied on a set of NMD genes that were originally identified in screens performed in yeast and *C. elegans* (Cali et al., 1999; Cui et al., 1995; Hodgkin et al., 1989; Leeds et al., 1991; 1992; Yamashita et al., 2009). A subsequent screen using RNAi-based methods in *C. elegans* identified two additional factors that appear to be required for NMD. These genes, called *smgl-1* and *smgl-2* (Longman et al., 2007; Rosains and Mango, 2012), are conserved in zebrafish (Anastasaki et al., 2011; Frizzell et al., 2012) and mammals (Cali et al., 1999; Longman et al., 2007), but their function in NMD is not well understood. Coimmunoprecipitation experiments with SMG-1 in HeLa cells identified SMG-8 and SMG-9 as protein interactors (Metzstein and Krasnow, 2006; Yamashita et al., 2009), which have been proposed to function in NMD. However, a conserved functional role for these proteins in NMD in intact organisms has not been demonstrated. Indeed, genetic analysis in *C. elegans* recently showed that *smg-8* does not contribute to NMD in that organism (Rosains and Mango, 2012; Xu and Rubin, 1993).

We have performed forward genetic screens for genes functionally involved in nonsense mediated decay in *Drosophila* (Frizzell et al., 2012). These screens are the first mutagen-based forward genetic screens since those performed in *C. elegans* that led to the identification of *smg-7* (Cali et al., 1999; Frizzell et al., 2012) and are also the first specifically designed to identify NMD genes in *Drosophila*. Our screening strategy utilizes an NMD-sensitive mRNA reporter (Frizzell et al., 2012; Metzstein and Krasnow, 2006). When this reporter is expressed in an NMD-deficient cell, fluorescent expression is drastically enhanced. We screened the X and right arm of the third chromosome (3R) using a mosaic approach, which allowed for the identification of both viable and lethal alleles of genes that are involved in NMD (Xu and Rubin, 1993). We identified multiple alleles on the X chromosome, all of which were alleles of the three previously identified NMD genes on this chromosome (Frizzell et al., 2012; Metzstein and Krasnow, 2006); therefore, the X chromosome is saturated for NMD genes identifiable by our methods. The screen on 3R resulted in the isolation of the first alleles of the known NMD gene *Smg6* in *Drosophila*, and are described in Chapter 2 and 4 of this thesis (Frizzell et al., 2012; Shiga et al., 1996). Seven lines on 3R complemented *Smg6* for NMD reporter enhancement and, since there are no other orthologs of previously identified NMD genes on this chromosome, these lines likely contain mutations in novel NMD genes. This chapter describes preliminary mapping and identification of potentially novel NMD genes.

Materials and methods

Fly stocks

D. melanogaster lines were raised at 25°C on cornmeal/dextrose medium using standard conditions. Complementation tests were performed using alleles described in (Frizzell et al., 2012), including *e22c-GAL4, UAS-nlsDsRed2::SV40 3'UTR, UAS-FLP/CyO; FRT^{82B} Smg6^{292}/TM6c Tb, Sb*. The control chromosome used was *FRT^{82B}* (Frizzell et al., 2012; Xu and Rubin, 1993). DrosDel deletions were obtained from the Bloomington *Drosophila* stock center and are listed in Table 3.2. The *UAS-eGFP:SV40 3'UTR* transgene was cloned as described in Chapter 5 and by (Metzstein and Krasnow, 2006; Venken et al., 2009). The *breathless* promoter *btl-Gal4* (Shiga et al., 1996; Venken et al., 2009) was used in combination with *UAS-GFP* as the reporter in the CG33970 RNAi tests. CG33970 RNAi lines *w^{1118}; P{GD297}v33088*, *w^{1118}; P{GD7619}v38660/CyO*, *w^{1118}; P{GD7619}v38661*, *w^{1118}; P{GD297}v46838/CyO*, and *P{KK109988}VIE-260B* and the Upf2 RNAi *w^{1118} P{GD9763}v33507* were obtained from the Vienna *Drosophila* RNAi center (Dietzl, 2007). *Act5c-GAL4* was acquired from the Bloomington *Drosophila* stock center.

Complementation tests

The X chromosome and right arm of the third chromosome (3R) were screened as described in (Frizzell et al., 2012; Venken et al., 2006). Complementation tests on 3R were performed by crossing *e22c-GAL4, UAS-nlsDsRed2::SV40 3'UTR, UAS-FLP/CyO; FRT^{82B}*/TM6c Tb, Sb* females (where * is the mutant being tested) to *Df(3R)ED6220/TM6c Tb, Sb,* or to *e22c-GAL4, UAS-nlsDsRed2::SV40 3'UTR, UAS-*

FLP/CyO; FRT^{82B} Smg6^{292}/TM6c Tb, Sb animals and scoring for DsRed enhancement in third-instar larvae. The deficiency *Df(3R)ED6220* removes approximately 644 kb of 3R, including the *Smg6* transcriptional unit along with ~115 other predicted genes. *Smg6* mutants, which fail to complement *Df(3R)ED6220* for enhanced *DsRed* expression, are described in (Frizzell et al., 2012; Rees et al., 2009). Seven lines on 3R complemented the NMD reporter enhancement when crossed to *e22c-GAL4, UAS-nlsDsRed2::SV40 3'UTR, UAS-FLP/CyO; FRT^{82B} Smg6^{292}/TM6c Tb, Sb* mutant animals (Fig. 3.1). To establish complementation groups, the lines were mated *inter se* and scored for loss of the *TM6c Tb, Sb* balancer chromosome (Fig. 3.1).

Lethal mapping

A minimal set of deficiencies from the DrosDel (ED) deficiency collection were identified that tile across most of the right arm of the third chromosome (Table 3.2). The deficiency stocks contain a balancer chromosome identifiable by the dominant bristle marker *Sb*. Complementation tests were performed by crossing these deficiencies to *FRT^{82B} 44b /TM6c Tb Sb*, a representative allele from the lethal complementation group, and scoring for failure to lose the balancer chromosome. A single deficiency, *Df(3R)ED6232*, was identified that failed to complement the lethality associated with *44b*. The other four lethal lines were crossed to *ED6232*, and were found to fail to complement for lethality. Additional deficiencies (Table 3.2) were then used to further narrow down the lethal region using the same methods.

Sequence analysis of lethal complementation group candidate genes

By deficiency mapping, the lethality associated with the lethal complementation group was mapped to a 181 kb region of 3R that contained twelve genes. Tiling primers were designed using Primer3 (http://frodo.wi.mit.edu) across the open reading frames of each of these genes (Table 3.2). For a template, each of the mutant lines was mated to the original FRT^{82B} isogenized chromosome. Individual nonbalancer adult flies were collected and DNA was isolated using a proteinase K protocol. The sequencing primers were used to amplify ~500 bp amplicons, which were visualized on a 1% agarose gel using an Alpha Innotech Gel Doc System. If a single band was present, the sample was treated with exo-SapIt (Affymetrix) and sequenced. For primers that resulted in multiple bands, the band of the correct size was excised from the agarose gel and purified using a gel purification kit (Qiagen), and then sequenced (Fig. 3.2).

Mosaic analysis of different fluorescent reporters in lethal group mutants

To determine if the lethal complementation group affected NMD function, I tested reporter enhancement using a different fluorescent reporter. For this assay *e22c-GAL4, UAS-nlsDsRed2::SV40 3'UTR, UAS-FLP/CyO;* FRT^{82B} *56c /TM6c Tb, Sb* or *e22c-GAL4, UAS-nlsDsRed2::SV40 3'UTR, UAS-FLP/CyO;* FRT^{82B} *44b /TM6c Tb, Sb* animals were mated to *UAS-eGFP::SV40 3'UTR* animals; Sb^{+} wandering third-instar larvae were collected that expressed both eGFP and DsRed reporters, heat fixed, and imaged using a Leica MZ16 fluorescence stereomicroscope and Leica DFC340 FX imager (Fig. 3.3 B,C). A *Smg6* mutant, *e22c-GAL4, UAS-nlsDsRed2:SV40 3'UTR, UAS-FLP/CyO;* FRT^{82B}

Smg6^{292}*/TM6c Tb, Sb,* was also mated and collected to show the eGFP fluorescence can be seen in a mosaic animal (Fig. 3.3 A).

RNAi experiments

To knock down CG33970 using RNAi, five RNAi lines were mated to *Btl-Gal4 UAS-eGFP::SV40 3'UTR* to test for NMD reporter enhancement. Fluorescent animals were collected in glycerol and examined on a Leica MZ16 fluorescence stereomicroscope. All five RNAi lines were also mated to *Act5c-GAL4* to test for adult lethality.

Genomic rescue of CG33970

To clone a genomic rescue construct of CG33970, two BAC clones, CH321-93I23 and CH321-44M01, were ordered from the BACPAC Genomic Resource Center (BPRC). These clones were from the CHORI-321 collection of 83 kb average clone sizes (Duffy et al., 1998; Venken et al., 2009) in order to encompass the large ~47 kb gene region of CG33970. The clones were grown according to the P[acman] protocol (Venken et al., 2009; Weaver and Krasnow, 2008) and plasmids were purified using the a large construct kit (Qiagen). The CH321-44M1 clones were sent for injection by Genetic Services Inc. Two injection sites were used, attp16 and attp40 (Venken et al., 2006; Weaver and Krasnow, 2008).

cDNA rescue construct of CG33970

A cDNA clone encompassing CG33970 was ordered from the DGRC, clone # R14039, and transformed into TOP10 cells. Plasmids were isolated using the QIAprep

spin miniprep kit (Qiagen) and the cDNA was amplified using primers with *NotI* and *StuI* restriction enzyme sites incorporated, CG33970cDNA L1 (GAAGGCGGCCGCATG) and CG33970cDNA R1 (GAAGAGGCCTTTCGGCAAAGTTTTTCTATC). The 2,759 bp amplicon was amplified using phusion high fidelity polymerase (NEB), digested using *NotI* and *StuI* (NEB), and ligated into *pUAST attB* (a gift from Jyoti Misra, Thummel lab). The resulting plasmid was then transformed into TOP10 cells (Invitrogen), sequenced, and restriction digested to verify the insert was correctly cloned into the vector. The plasmid was then isolated using the plasmid midi kit (Qiagen) and prepared for injection using Genetic Services Inc.

Imaging nonlethal complementation group alleles

Two alleles on 3R complemented both *Smg6* and the lethal complementation group (Fig. 3.1). Mosaics of these alleles were made by mating *e22c-GAL4, UAS-nlsDsRed2::SV40 3'UTR, UAS-FLP/CyO; FRT^{82B} 49c* or *92c /TM6c Tb, Sb* males to *FRT^{82B}* females. Wandering third-instar larvae that had mosaic expression of DsRed were collected, heat fixed, and imaged using a Leica MZ16 fluorescence stereomicroscope and Leica DFC340 FX imager (Fig. 3.4 A,B).

Results

Mutations in the lethal complementation group map to the novel gene, CG33970

To determine the identity of the gene causing lethality in the lethal complementation group, coding regions of 12 genes uncovered by deficiencies on 3R

were sequenced, revealing missense mutations in a single gene, CG33970 (Fig. 3.2 B). CG33970 is a predicted ABC transporter consisting of six transmembrane domains and an ATP binding cassette (Rees et al., 2009; Venken and Bellen, 2005; Venken et al., 2006; 2009). The lethal complementation group consists of five lines, four of which, *44b*, *44c*, *43d*, and *50a,* encode the same missense mutation. This mutation, C1883T, converts a serine to a phenylalanine at a moderately conserved residue in a predicted transmembrane domain (Fig. 3.2 B,C). A single line, *56c*, contains a distinct mutation in the same gene, G619A, which converts an aspartic acid to an asparagine in a highly conserved residue of a Walker domain in the ATP binding cassette. The mutation in *56c* is predicted to prevent ABC-transporter function (Fig 3.2 B,C).

Mutations in CG33970 may not cause an NMD-defect

One problem with the analysis of the lethal complementation group is that mapping was based on viability (Fig. 3.1), not complementation of an NMD defect. Therefore, it is not clear if the lethality associated with this complementation group is also responsible for the NMD defect. I next tested if the lethal lines exhibit a hemizygous NMD defect that was detectable earlier in development. A mutant line carrying the reporter was placed in *trans* to the deficiency that failed to complement lethality. The resulting embryos were examined for enhanced reporter fluorescence (data not shown), which would indicate the NMD defect was associated with the lethality-mapped deficiency region. However, mutant/Df lines exhibit embryonic lethality and while *e22c-GAL4* drives expression ubiquitously in embryos (Duffy et al., 1998; Venken et al., 2006; 2009), the early lethality does not provide enough time for DsRed to mature, and no

fluorescent signal could be detected (Dean and Annilo, 2005; Weaver and Krasnow, 2008). These two difficulties made it impossible to determine if the NMD defect and lethal defect were linked using this assay. In future work, this experiment could be done using a GFP-based NMD assay, as the maturation time of GFP is significantly less than that of DsRed (Rees et al., 2009; Weaver and Krasnow, 2008).

Another explanation for these lethal alleles not exhibiting an NMD defect is that the mosaic cells have a modestly different phenotype than other NMD mosaics. First, the nuclei appear smaller than neighboring nuclei. Second, mutant clones show very weak, albeit reproducible, reporter enhancement; I was able to positively identify them in several retests. To verify the mosaic enhancement in the lethal complementation group is indeed an NMD defect and not a concentration of the DsRed protein caused by a condensed nucleus, the mutants were crossed to a different set of NMD reporters, which utilize the fluorescent transgene GFP instead of DsRed. This GFP transgene has been used before to identify NMD-mutant mosaic cells. As a positive control, I tested this transgene in a *Smg6* mutant and was able to positively identify enhanced nuclear localized DsRed as well as enhanced cytoplasmically localized GFP in the same mosaic cells (Fig 3.3A). When this test was performed in two lethal lines, *44b* and *56c*, the DsRed enhancement was weak, but identifiable. However, the GFP showed no clear enhancement (Fig. 3.3 B,C).

RNAi knockdown of CG33970 does not cause NMD defects

To test if loss of CG33970 is sufficient to cause enhancement of an NMD-sensitive reporter, RNAi knock down of CG33970 was performed. Five RNAi lines

directed against CG33970 were mated to *Btl-Gal4 UAS-eGFP::SV40 3'UTR* to determine if knockdown of CG33970 in the tracheal system results in enhancement of the GFP reporter. I found none of these lines resulted in a visible enhancement of the GFP reporter (data not shown). As a control, *Upf2* RNAi was also crossed to *Btl-Gal4 UAS-eGFP:SV40 3'UTR*, which resulted in increased GFP fluorescence in the trachea (data not shown). As a second control, each of the CG33970 RNAi lines was crossed to *Act5c-GAL4* to test if ubiquitous CG33970 knockdown causes lethality, as would be expected based on the recovered alleles. Of the five lines, only one, w^{1118}; *P{GD7619}v38660/CyO*, was lethal.

CG33970 rescue

To test if mutations in CG33970 indeed cause lethality in these lines, I made two attempts at cloning rescue constructs for CG33970. First, I attempted to make a genomic rescue construct of CG33970. To clone a rescuing construct I utilized the BACPAC clonal library, which consists of large, 83 kb on average, BAC-clones that contain the sequence. This allowed for site specific integration using the P[acman] cloning and integration system (Dean et al., 2001; Venken and Bellen, 2005; Venken et al., 2006; 2009). Two injection sites were attempted with multiple injections; however, neither injection site resulted in the incorporation of a stable construct. Failure to incorporate into the genome was verified by the lack of w^{+} transgene that marks the clone and lack of detectable transgene insertion via PCR of the BAC clone ends (data not shown). Attempts at rescue using any animals with portions of these constructs resulted in whole animal lethality. The failure to isolate lines of this construct could have resulted for

several reasons: first, the efficiency of large >50 kb clones is low using this technique (2-4% transformation efficiency) (Venken et al., 2006; 2009). To overcome this obstacle we injected clones multiple times in different locations in the genome, but were unsuccessful in isolating a stable genomic rescue line. Second, the presence of extra copies of CG33970 may not be tolerated for viability. To avoid this possibility, I made a UAS-CG33970 expression vector. Perhaps this smaller vector encoding only the cDNA of CG33970 will provide an answer as to whether lethality of CG33970 can be rescued by expression of this gene. This construct must still be injected and stable lines must be made to test whether CG33970 is the gene responsible for lethality in this interval. Additionally, more cloning attempts and integration sites could be tried to rescue CG33970. Another route could be to test whether CG33970 is directly involved in NMD in S2 cells. If so, this is an interesting candidate to continue rescue attempts, and if not, this may be a gene that needs no further analysis in our laboratory.

Nonlethal complementation group genes

Out of the seven lines isolated on 3R that complemented *Smg6*, two lines complemented the lethality associated with the lethal complementation group. These lines are called *92c* and *49c*. These lines also complement each other for lethality and for reporter enhancement (Fig. 3.1), indicating they represent two potentially novel NMD genes. *49c* displays a weak mosaic phenotype (Fig. 3.4), while *92c* displays a variable weak to strong mosaic phenotype (Fig. 3.4). These alleles occasionally have adult escapers, however, the frequency has not been quantified. Mapping of these alleles should be performed using the minimal deficiency collection (Table 3.1) and would be

especially efficient if a fluorescent reporter were established in the mutant lines to test for both failure to complement lethality and NMD reporter enhancement.

Discussion

Mapping lethality and sequence analysis have thus far resulted in the identification of CG33970, a previously uncharacterized gene in *Drosophila*, whose potential role in NMD remains elusive. Sequence and homology analysis suggest CG33970 belongs to the ABC transporter superfamily (Dean and Annilo, 2005). ABC transporters are an extensive family of transmembrane proteins that translocate a variety of substances across both intra- and extra-cellular membranes (Rees et al., 2009). CG33970 belongs to a group, denoted ABCH, that contains only two other *Drosophila* genes. Both of these are also uncharacterized and their domains more closely resemble prokaryotic domains than eukaryotic ABC domains (Dean et al., 2001). Protein homology suggests CG33980 is conserved in invertebrates, but is not found in vertebrate lineages. Intriguingly, the most closely related transporter in *E. coli* is a ribosome bound ABC transporter called RbbA, encoded by the transcriptional unit *yhih* (Cali et al., 1999; Cui et al., 1995; Dean and Annilo, 2005; Hodgkin et al., 1989; Kiel and Ganoza, 2001; Kiel et al., 1999; Leeds et al., 1991; 1992), which may function in ribosome translation termination (Longman et al., 2007; Xu et al., 2006). A possible conservation of function between these two genes raises the possibility that CG33970 could be involved in an early step of NMD, potentially during translation termination and recognition of the premature termination codon. If this were the case, the weak NMD defect might be due to compensation by other standard factors required for translation termination. This

could also be a transcript-specific mechanism, which might explain the lack of enhanced GFP in mosaic epidermal cells, while DsRed is enhanced. Another interesting possibility is the transmembrane domains of CG33970 could potentially link NMD to a specific organelle, like the endoplasmic reticulum, to allow for more localized regulation of RNA metabolism. Localization of NMD components with the endoplasmic reticulum has been observed in *C. elegans* and HeLa cells (Anastasaki et al., 2011; Sakaki et al., 2012); thus, the discovery of CG33970 could represent an NMD factor that localizes NMD function to a particular organelle, potentially during times of stress. Local regulation of translation and decay is a common feature of RNA control, indeed, this has been observed in neurons with NMD-targeted transcripts (Long et al., 2010; Longman et al., 2007; Moore, 2005).

Identification of mutations in a gene is not sufficient evidence that the particular gene plays a role in that process. The gene should be capable of rescuing the mutant phenotype, and if this is not possible, a secondary loss of function method provides more evidence for the role of the gene in a given process. In this case, several lines of evidence suggest CG33970 may not be involved in NMD.

While two independent mutations were identified in CG33970, four lines contained the same coding change. There are three possible explanations that may explain the presence of the same mutation in four lines. First, the mutation could occur in a highly mutable region of the genome. Second, the four lines may have been duplicated from a single isolated mutation. A duplication of a line would be the result of human error, which, while possible, seems highly unlikely as two of the lines were isolated during independent rounds of screening. Third, the mutations in CG33970 could be

SNPs present in the background of the mutagenized stocks. This is unlikely as the parental stocks were isogenized and sequencing parental chromosomes did not reveal SNPs in CG33970 (data not shown). However, not enough animals (one representative of each male and female) were sequenced to make a population level conclusion. The presence of a SNP in CG33970 still could also be the result of the way these mutations were mapped, as complementation tests mapped lethality associated with these lines and assumed that the NMD defect was lethal. There is a chance that the lethality of these lines is not associated with the putative NMD defect, resulting in the identification of an extraneous lethal mutation

Another line of evidence that these alleles may not actually disrupt NMD function is that they exhibit weak *nlsDsRed2* reporter stabilization in mosaics when compared to other NMD mutants. When the lethal mutants were tested for stabilization of a different fluorescent reporter, no stabilization in homozygous mutant cells was observed. The lack of stabilization of the GFP reporter in mosaic cells (recognized by DsRed enhancement) could have been caused by the diffuse localization of the GFP reporter, which may have prohibited detection of reporter stabilization using this technique. One way to address whether CG33970 has an NMD defect would be to use a nuclear GFP; however, this will not differentiate between a defect in reporter accumulation due to other defects, like nuclear export, therefore, other methods to validate the role if this gene in NMD must be used. These data support a non-NMD role for the defect in CG33970, as the cytoplasmic GFP reporter is clearly stabilized in other NMD mutants, such as *Smg6*.

Another way to test if CG33970 is involved in NMD is RNAi knockdown of the transcript. Five independent RNAi lines were crossed to NMD reporter lines, none of

which caused any detectable visible defects, however, more controls must be done to verify this result. First, knock down of CG33970 needs do be verified using qRT-PCR. Expressing these RNAi lines using a ubiquitous driver was expected to cause lethality due to the mutant phenotype; however, only one of the five lines was lethal, suggesting there may be inefficient knockdown of CG33970. In future experiments a variety of ubiquitous GAL4 drivers should be used and the knockdown should be verified using qRT-PCR. In addition, the NMD defect should be re-tested using *e22c-GAL4 UAS-nlsDsRed:SV40 3'UTR*, since this is the reporter originally used to identified the alleles of CG33970. Mosaic cells were not identifiable using the GFP reporter that was used in this experiment. This suggests this defect is either fluorescent reporter specific, and therefore not an NMD defect, or minor changes in expression are difficult to see using GFP in these alleles. To detect minor perturbations in reporter expression, qRT-PCR against GFP and DsRed could be performed.

One method that could be utilized to determine if CG33970 is involved in NMD is to knock down CG33970 in *Drosophila* S2 cells. NMD functions in S2 cells and if CG33970 is also expressed in this cell type, knock down followed by qRT-PCR of known NMD targets will provide a direct assay for whether CG33970 plays a role in NMD.

In addition to the lethal complementation group, there are two lines that likely represent new NMD genes, *92c* and *49c*. These lines have yet to be mapped or sequenced and will be an interesting new group to study. The first thing that should be done with these lines is rebalancing an NMD reporter in the background to provide a

readout of NMD function; this will aid in deficiency mapping that can be performed to narrow down the region of the mutations.

The identification of new genes involved in a particular process is a primary goal of performing a forward genetic screen. Using our reporter-based screening strategy we have potentially identified new genes involved in NMD. There is, however, a large amount of work that still needs to be done to map, identify, verify, and finally characterize these genes and their role in NMD. The low number of alleles, one of each in the last two complementation groups, suggest 3R is not yet saturated for alleles of genes identifiable by our reporter, so further screening to saturation could be fruitful. Characterizing lethal mutations is a difficult task. The gene associated with the lethal complementation group, CG33970, has proven difficult to work with as cloning and rescue have not been possible, potentially due to the sheer size of the gene (47 kb). Rescue of both lethality and the NMD defect must be done to prove that CG33970 is the causative gene in this situation. This may be challenging, as other genes, including *Upf1*, are sometimes difficult to rescue, presumably due to enhanced regulation of gene expression. However, the cloning must first be accomplished to attempt further experiments. In the future more screening should be performed to saturate the *Drosophila* genome for alleles of NMD genes. Any tools we can establish in *Drosophila* will greatly further our understanding of NMD *in vivo*.

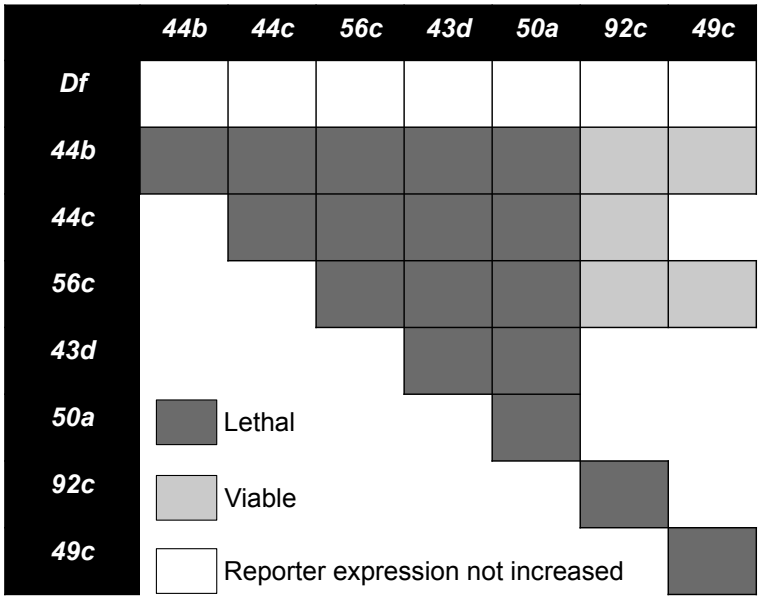


Figure 3.1 Complementation tests for non-*Smg6* 3rd chromosome alleles. When placed in *trans* to a deficiency uncovering *Smg6,* seven lines complement the deficiency for NMD-reporter enhancement (white boxes) and viability. Five of the lines are homozygous lethal, and are lethal as *trans*-heterozygotes (dark grey boxes). Two of the lines are lethal as homozygotes (with some escapers), and are viable as *trans*-heterozygotes with each other or with representative alleles from the lethal complementation group (light grey boxes).

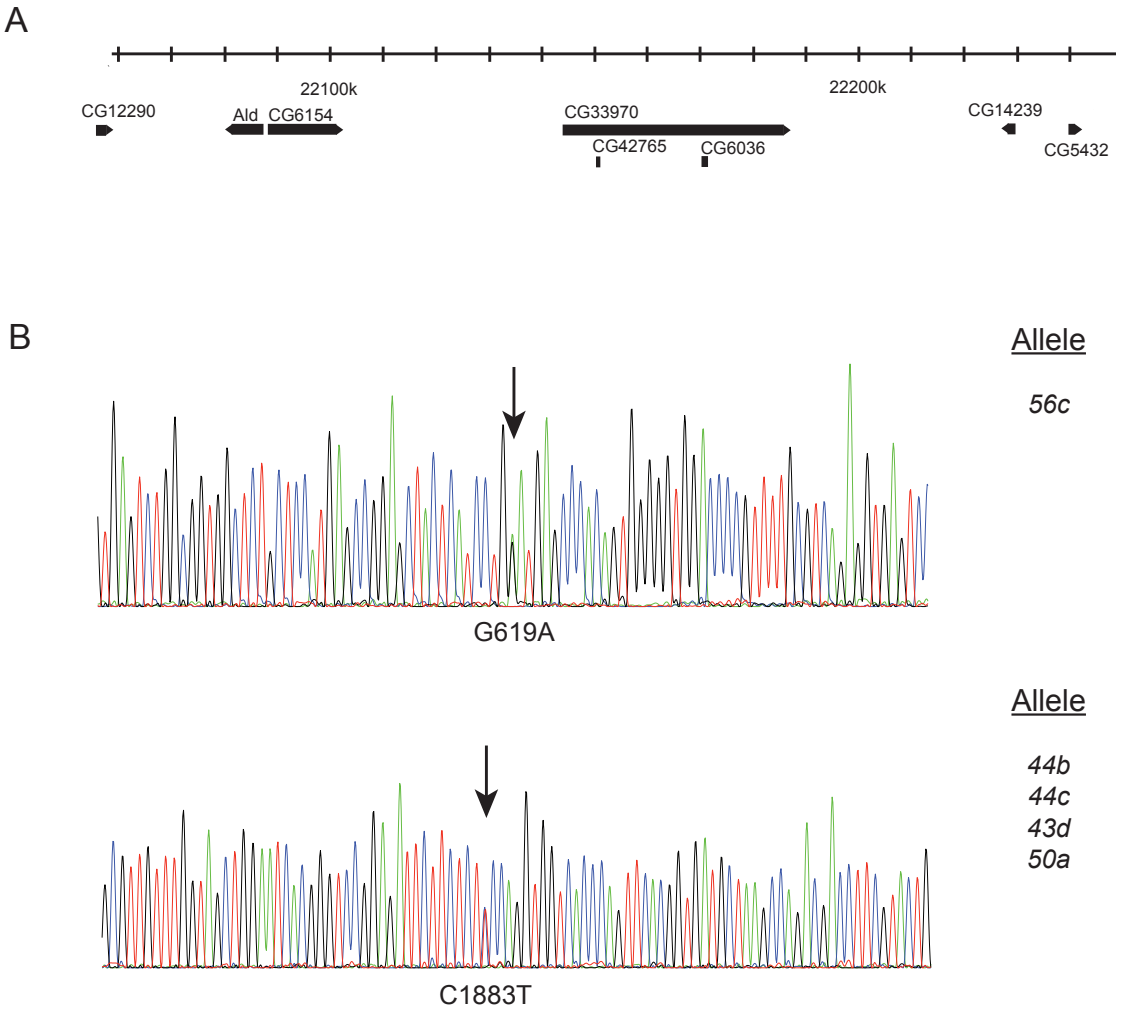


Figure 3.2 Identification of lethal mutations. (A) 181 kb region mapped by deficiencies that failed to complement the lethal complementation group. (B) Sanger sequencing of amplicons in CG33970 in two representative alleles from the lethal complementation group. Allele names listed on the right are alleles that contained the same mutation as the sequenced allele. (C) Predicted CG33970 protein structure and locations of the mutations within the ATPase and transmembrane domains.

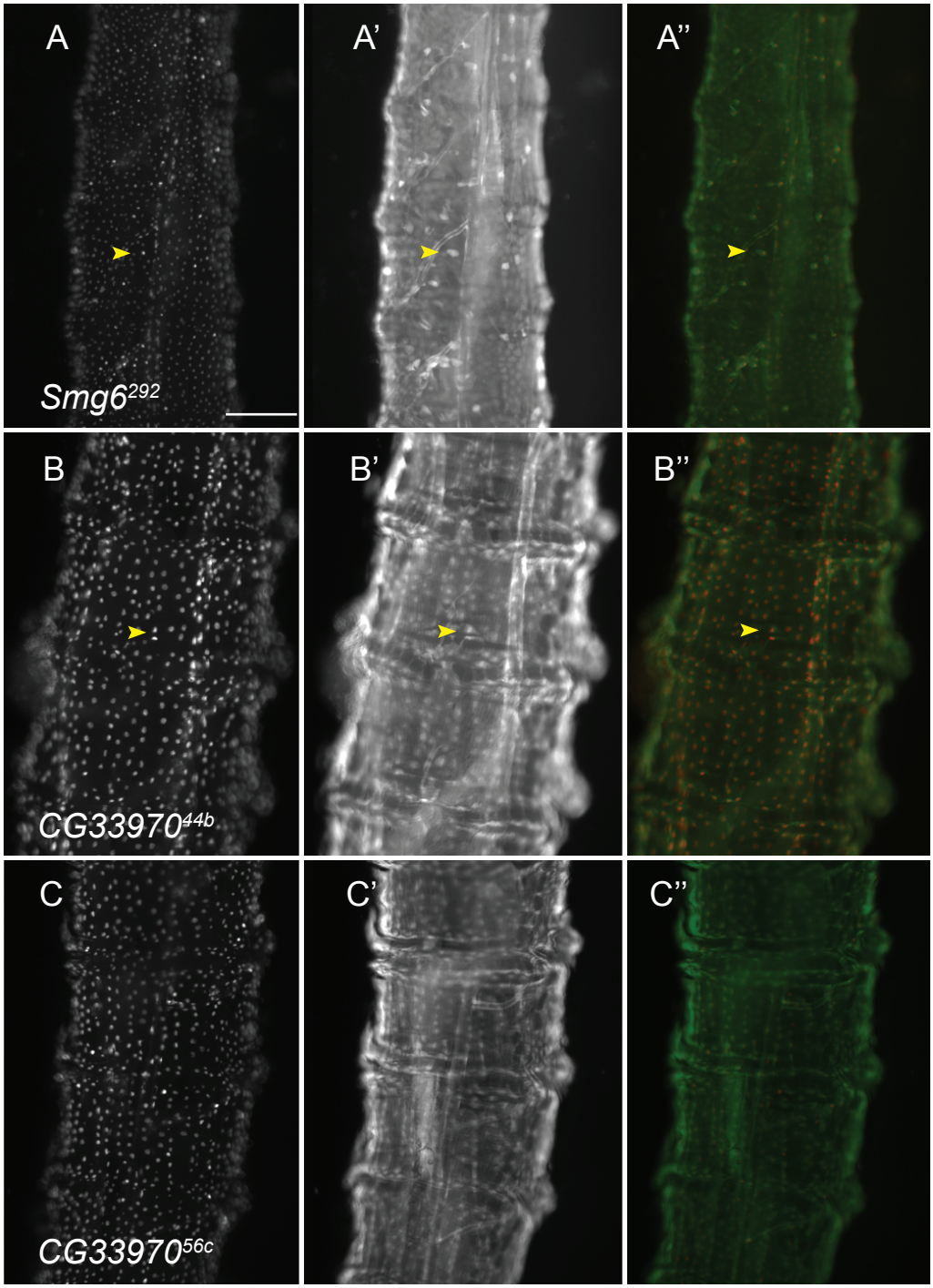


Figure 3.3 Mosaic analysis of NMD reporters in lethal mutants. (A-C) Red channel, (A'-C') Green channel, (A''-C'') merge. (A-A'') *e22c-GAL4, UAS-nlsDsRed2::SV40 3'UTR, UAS-FLP/ UAS-eGFP::SV40 3'UTR; FRT*^{82B} *Smg6*^{292} /+. (B-B'') *e22c-GAL4, UAS-nlsDsRed2::SV40 3'UTR, UAS-FLP/ UAS-eGFP::SV40 3'UTR; FRT*^{82B} *CG3397*^{44b} /+. (C-C'') *e22c-GAL4, UAS-nlsDsRed2::SV40 3'UTR, UAS-FLP/ UAS-eGFP::SV40 3'UTR; FRT*^{82B} *CG33970*^{56c} /+. Yellow arrowheads indicate enhanced mosaics. Scale bar 0.5mm.

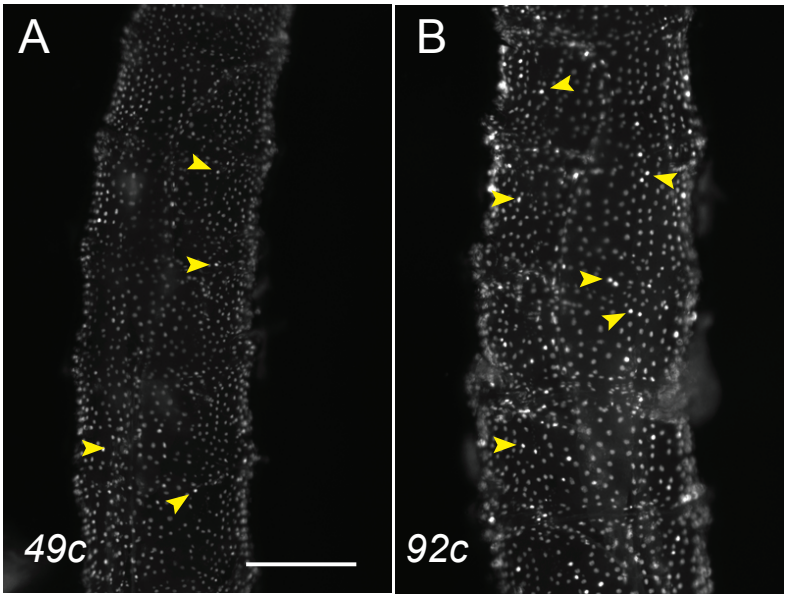


Figure 3.4 Mosaics in additional alleles. (A) *e22c-GAL4, UAS-nlsDsRed2::SV40 3'UTR, UAS-FLP/ +; FRT^{82B} 49c / FRT^{82B}*. (B) *e22c-GAL4, UAS-nlsDsRed2::SV40 3'UTR, UAS-FLP/ +; FRT^{82B} 92c / FRT^{82B}*. Yellow arrowheads indicate a subset of enhanced nuclei. Error bar 0.5mm.

Table 3.1 3R deficiencies used to map lethal mutations.

Bloomington Stock #	Genotype	Breakpoints	Viability
6962	w[1118]; Df(3R)ED2, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}CG 31232[ED2]/TM6C, cu[1] Sb[1]	91A5;91F1 (R4 estimated cytology), 3R:14224953..14922493 (R4)	Viable
8029	w[1118]; Df(3R)ED5577, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 5577/TM6C, cu[1] Sb[1]	86F9;87B13 (R4 estimated cytology), 3R:7654463..8303300 (R4)	Viable
8091	w[1118]; Df(3R)ED5092, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 5092/TM6C, cu[1] Sb[1]	82A1;82E7 (R4 estimated cytology), 3R:107408..912807 (R4)	Viable
8103	w[1118]; Df(3R)ED5177, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 5177/TM6C, cu[1] Sb[1]	83B4;83B6 (R4 estimated cytology), 3R:1426351..1449817 (R4)	Viable
8104	w[1118]; Df(3R)ED5780, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 5780/TM2	89E11;90C1 (R4 estimated cytology), 3R:12882199..13507523 (R4)	Viable
8105	w[1118]; Df(3R)ED6232, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 6232/TM6C, cu[1] Sb[1]	96F10;97D2 (R4 estimated cytology), 3R:21862598..22624704 (R4)	Lethal
8680	w[1118]; Df(3R)ED5138, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 5138/TM6C, cu[1] Sb[1]	82D5;82F8 (R4 estimated cytology), 3R:606794..1090605 (R4)	Viable
8682	w[1118]; Df(3R)ED5230, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 5230/TM6C, cu[1] Sb[1]	84E6;85A5 (R4 estimated cytology), 3R:3803496..4478856 (R4)	Viable
8683	w[1118]; Df(3R)ED5911, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 5911/TM6C, cu[1] Sb[1]	91C5;91F4 (R4 estimated cytology), 3R:14568649..14991505 (R4)	Viable
8684	w[1118]; Df(3R)ED6096, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 6096/TM6C, cu[1] Sb[1]	94B5;94E7 (R4 estimated cytology), 3R:18413403..19047691 (R4)	Viable
8685	w[1118]; Df(3R)ED7665, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 7665/TM6C, cu[1] Sb[1]	84B4;84E11 (R4 estimated cytology), 3R:2916249..3919805 (R4)	Viable
8922	w[1118]; Df(3R)ED5942, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 5942/TM6C, cu[1] Sb[1]	91F12;92B3 (R4 estimated cytology), 3R:15052016..15660809 (R4)	Viable
8925	w[1118]; Df(3R)ED6316, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 6316/TM6C, cu[1] Sb[1]	99A5;99C1 (R4 estimated cytology), 3R:25081045..25608389 (R4)	Viable
8957	w[1118]; Df(3R)ED5514, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 5514/TM6C, cu[1] Sb[1]	86C7;86E11 (R4 estimated cytology), 3R:6710720..7394975 (R4)	Viable
8960	w[1118]; Df(3R)ED6265, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 6265/TM6C, cu[1] Sb[1]	97E2;98A7 (R4 estimated cytology), 3R:22937981..23405492 (R4)	Viable
8961	w[1118]; Df(3R)ED6310, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 6310/TM6C, cu[1] Sb[1]	98F12;99B2 (R4 estimated cytology), 3R:24964617..25337875 (R4)	Viable
8962	w[1118]; Df(3R)ED6076, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED 6076/TM6C, cu[1] Sb[1]	93E10;94A1 (R4 estimated cytology), 3R:17459227..17868550 (R4)	Viable

Table 3.1 Continued

Bloomington Stock #	Genotype	Breakpoints	Viability
8965	w[1118]; Df(3R)ED5156, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED5156/TM6C, cu[1] Sb[1]	82F8;83A4 (R4 estimated cytology), 3R:1090655..1284574 (R4)	Viable
8967	w[1118]; Df(3R)ED5147, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED5147/TM6C, cu[1] Sb[1]	82E7;83A1 (R4 estimated cytology), 3R:912842..1193526 (R4)	Viable
9082	w[1118]; Df(3R)ED5474, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED5474/TM6C, cu[1] Sb[1]	85F11;86B1 (R4 estimated cytology), 3R:5935134..6176446 (R4)	Viable
9084	w[1118]; Df(3R)ED5518, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED5518/TM6C, cu[1] Sb[1]	86C7;86E13 (R4 estimated cytology), 3R:6710720..7445622 (R4)	Viable
9086	w[1118]; Df(3R)ED5591, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED5591/TM6C, cu[1] Sb[1]	87B7;87C7 (R4 estimated cytology), 3R:8176253..8545732 (R4)	Viable
9087	w[1118]; Df(3R)ED5610, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED5610/TM6C, cu[1] Sb[1]	87B11;87D7 (R4 estimated cytology), 3R:8269738..8821397 (R4)	Viable
9088	w[1118]; Df(3R)ED5612, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED5612/TM6C, cu[1] Sb[1]	87C7;87F6 (R4 estimated cytology), 3R:8545707..9470856 (R4)	Viable
9090	w[1118]; Df(3R)ED5644, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED5644/TM6C, cu[1] Sb[1]	88A4;88C9 (R4 estimated cytology), 3R:9843625..10451431 (R4)	Viable
9159	w[1118]; Df(3R)ED10257, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED10257/TM6C, cu[1] Sb[1]	83A7;83B4 (R4 estimated cytology), 3R:1344078..1426000 (R4)	Viable
9198	w[1118]; Df(3R)ED5142, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED5142/TM6C, cu[1] Sb[1]	82B2;82F8 (R4 estimated cytology), 3R:279018..1090605 (R4)	Viable
9204	w[1118]; Df(3R)ED5339, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED5339/TM6C, cu[1] Sb[1]	85D1;85D11 (R4 estimated cytology), 3R:5052798..5178097 (R4)	Viable
9209	w[1118]; Df(3R)ED6025, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED6025/TM6C, cu[1] Sb[1]	92A11;92E2 (R4 estimated cytology), 3R:15468450..16135241 (R4)	Viable
9210	w[1118]; Df(3R)ED6255, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED6255/TM6C, cu[1] Sb[1]	97D2;97F1 (R4 estimated cytology), 3R:22624758..23107623 (R4)	Viable
9211	**w[1118]; Df(3R)ED6220, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED6220/TM6C, cu[1] Sb[1]**	**96A7;96C3 (R4 estimated cytology), 3R:20369520..21009495 (R4)**	**Viable**
9215	w[1118]; Df(3R)ED5495, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED5495/TM6C, cu[1] Sb[1]	85F16;86C7 (R4 estimated cytology), 3R:5996223..6712482 (R4)	Viable
9229	w[1118]; Df(3R)ED5797, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED5797/TM6C, cu[1] Sb[1]	90C2;90F10 (R4 estimated cytology), 3R:13543832..14068391 (R4)	Viable
9338	w[1118]; Df(3R)ED5296, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED5296/TM6C, cu[1] Sb[1]	84F6;85C3 (R4 estimated cytology), 3R:4076143..4882413 (R4)	Viable

Table 3.1 Continued

Bloomington Stock #	Genotype	Breakpoints	Viability
9478	w[1118]; Df(3R)ED6235, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED6235/TM6C, cu[1] Sb[1]	97B9;97D12 (R4 estimated cytology), 3R:22360956..22806229 (R4)	Viable
9481	w[1118]; Df(3R)ED10639, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED10639/TM6C, cu[1] Sb[1]	89B7;89D2 (R4 estimated cytology), 3R:12038635..12306942 (R4)	Viable
9482	w[1118]; Df(3R)ED10642, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED10642/TM2	89C7;89D5 (R4 estimated cytology), 3R:12279479..12450993 (R4)	Viable
9483	w[1118]; Df(3R)ED10809, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED10809/TM6C, cu[1] Sb[1]	92F1;93B8 (R4 estimated cytology), 3R:16376838..16886270 (R4)	Viable
9486	w[1118]; Df(3R)ED10820, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED10820/TM6C, cu[1] Sb[1]	93A4;93B12 (R4 estimated cytology), 3R:16774462..16937182 (R4)	Viable
Cambridge	w[1118]; Df(3R)ED6058/TM6C, cu[1] Sb[1]	93D4; 93F6	Viable
Cambridge	w[1118]; Df(3R)ED6105/TM6C, cu[1] Sb[1]	94E9; 94E11	Viable
Cambridge	w[1118]; Df(3R)ED6290/TM6C, cu[1] Sb[1]	98C3; 98E5	Viable
Cambridge	w[1118]; Df(3R)ED6332/TM6C, cu[1] Sb[1]	99E4; 99F2	Viable
Cambridge	w[1118]; Df(3R)ED6346/TM6C, cu[1] Sb[1]	100A5; 100B1	Viable
Cambridge	w[1118]; Df(3R)ED6361/TM6C, cu[1] Sb[1]	100C7; 100E3	Viable
Europe/FBab0044431	Df (3R) ED 10894	95A7;95E1	Viable
Europe/FBab0037564	Df (3L) ED 6346	100A5;100B1	Viable
Europe/FBab0044486	Df (3R) ED 10951	96B20;96D1	Viable
Europe/FBab0036895	Df (3R) ED 5660	88D1;88E5	Viable
24516	w[1118]; Df(3R)ED50003, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED50003/TM6C, cu[1] Sb[1]	100E1;100F5	Viable
24136	w[1118]; Df(3R)ED5416, P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED5416/TM6C, cu[1] Sb[1]	85D16;85E6	Viable
5601*	Df(3R)Espl3/TM6C, cu[1] Sb[1] Tb[1] ca[1]	96F1;97B1	Viable
24999*	w[1118]; Df(3R)BSC495/TM6C, Sb[1] cu[1]	96F6;97B4, 3R:21716131;22287273 (R5 flank)	Lethal
25016*	w[1118]; Df(3R)BSC512/TM6C, Sb[1] cu[1]	97A1;97B2, 3R:21932613;22265924 (R5 flank)	Lethal
26850*	w[1118]; Df(3R)BSC752/TM6C, Sb[1] cu[1]	97A2;97C5, 3R:22005377;22551031 (R5 flank;R3 author statement->R5)	Lethal
25000*	w[1118]; Df(3R)BSC496/TM6C, Sb[1] cu[1]	97A6;97D4, 3R:22084363--22084371;22702382 (R5 flank;R3 author statement->R5)	Lethal

Grey shading: deficiency that fails to complement lethal mutants.
Bold: deficiency that uncovers *Smg6*.
* Deficiencies used to narrow down the lethal mutant region.

Table 3.2: Primers used for sequencing candidate ORFs.

Primer Name	Primer Sequence
CG33970 L1	Gcgtaatgcgtacgtctgtatc
CG33970 L2	TCCAGGCAAAAGTAGGAGGA
CG33970 L3	TGCTGGTCGTGATACAGCTT
CG33970 L4	tatttgatcggcatgcacac
CG33970 L5	Gctcaactaaccgtgtgcaa
CG33970 L6	CGTCAGAGgtgagtccagaa
CG33970 L7	caagtcgcatcaaagtcgaa
CG33970 L8	GGCTCGAGATCCGATCAGTA
CG33970 L9	ggcaggtgcaatgacttttt
CG33970 L10	CAACTTTGTGGGTGACATGC
CG33970 L11	TGGAAATTGTTGTGCCTTGT
CG33970 L12	ccgttctggcttgacgata
CG33970 L13	cgaagagctgtgccaatttt
CG33970 L14	atgcagttccaatgccagtt
CG33970 L15	TCTGGCGAGATTTGGGTACT
CG33970 L16	tggagtcataaaccagcgaag
CG12290 L1	AACAGCAACTGCAGCAACAC
CG12290 L2	GAGCCAAAACAAAACCGAAA
CG12290 L3	GTCATCACGGAAACCGAGAC
CG12290 L4	CTGCAGAGCGCCTTGAAT
CG12290 L5	GCTCAGTGAGGAGTGTGGTTC
CG12290 L6	CAGAGTTTGCCCACTGAGGT
CG6154 L1	Caaacaagctggctttcaga
CG6154 L2	ATCGACGGACACAACGATCT
CG6154 L3	GACTGCTGCAAAGTTGACGA
CG6154 L4	ggtggtggtggtgcactatt
CG6154 L5	tcgccagaatgtttaccataac
CG6036 L1	TTCCGCACCTTGTACTACCC

Primer Name	Primer Sequence
CG33970 R1	TGACCTGCCGATGATTGTTA
CG33970 R2	AAATACGAGGCTGGCATACG
CG33970 R3	tgtgcttatggactcgatgc
CG33970 R4	GCCGAGTACCCAAATCTCG
CG33970 R5	tccactccactattcccaaca
CG33970 R6	aaaggacacggatttgttgc
CG33970 R7	GAAGGATGTGGGTATTTCCTG
CG33970 R8	cactcacCACAACCATGCTC
CG33970 R9	GGGATTGCGATAGCCATAAA
CG33970 R10	GATTCCGTTGGTTTGGTGAG
CG33970 R11	ACATGCATTTCCGGCTTAAT
CG33970 R12	tgcaaatagcccataaatcg
CG33970 R13	ccgataagctgtcaggcatc
CG33970 R14	ttccaggtggaaaatggaag
CG33970 R15	AACTCGCCGTAAAGGGCTAT
CG33970 R16	CCAGATGctgcaatcaaaaa
CG12290 R1	GCTGTTTTTCGATTACCACACA
CG12290 R2	CGCAGCAGTCATGTCTATGG
CG12290 R3	GCCGATGAACACACAAGCTA
CG12290 R4	GATGCTGTTGCTGCTGATGT
CG12290 R5	CAATTGTGCACCACTTGCAG
CG12290 R6	TGCGTCTCTTCTTGTTTAGCC
CG6154 R1	GCAGGTCTGTGTGACTCCAG
CG6154 R2	aacatctcacCTTGCCGAAC
CG6154 R3	ATGTGGTTGATATGCGctga
CG6154 R4	ccagccaaagtctggttagc
CG6154 R5	CTCACGCACAGAGCAACAAT
CG6036 L2	ACGCCAAAAATGGTTCGATA

References

Anastasaki, C., Longman, D., Capper, A., Patton, E.E., and Cáceres, J.F. (2011). Dhx34 and Nbas function in the NMD pathway and are required for embryonic development in zebrafish. Nucl. Acids Res. *39*, 3686–3694.

Cali, B.M., Kuchma, S.L., Latham, J., and Anderson, P. (1999). smg-7 is required for mRNA surveillance in Caenorhabditis elegans. Genetics *151*, 605–616.

Cui, Y., Hagan, K.W., Zhang, S., and Peltz, S.W. (1995). Identification and characterization of genes that are required for the accelerated degradation of mRNAs containing a premature translational termination codon. Genes Dev. *9*, 423–436.

Dean, M., Rzhetsky, A., and Allikmets, R. (2001). The human ATP-binding cassette (ABC) transporter superfamily. Genome Research *11*, 1156–1166.

Dean, M., and Annilo, T. (2005). Evolution of the ATP-binding cassette (ABC) transporter superfamily in vertebrates. Annu. Rev. Genomics Hum. Genet. *6*, 123–142.

Duffy, J.B., Harrison, D.A., and Perrimon, N. (1998). Identifying loci required for follicular patterning using directed mosaics. Development *125*, 2263–2271.

Frizzell, K.A., Rynearson, S.G., and Metzstein, M.M. (2012). Drosophila mutants show NMD pathway activity is reduced, but not eliminated, in the absence of Smg6. Rna *18*, 1475–1486.

Hodgkin, J., Papp, A., Pulak, R., Ambros, V., and Anderson, P. (1989). A new kind of informational suppression in the nematode Caenorhabditis elegans. Genetics *123*, 301–313.

Kiel, M.C., and Ganoza, M.C. (2001). Functional interactions of an Escherichia coli ribosomal ATPase. Eur. J. Biochem. *268*, 278–286.

Kiel, M.C., Aoki, H., and Ganoza, M.C. (1999). Identification of a ribosomal ATPase in Escherichia coli cells. Biochimie *81*, 1097–1108.

Leeds, P., Peltz, S.W., Jacobson, A., and Culbertson, M.R. (1991). The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon. Genes Dev. *5*, 2303–2314.

Leeds, P., Wood, J.M., Lee, B.S., and Culbertson, M.R. (1992). Gene products that promote mRNA turnover in Saccharomyces cerevisiae. Mol. Cell. Biol. *12*, 2165–2177.

Long, A.A., Mahapatra, C.T., Woodruff, E.A., Rohrbough, J., Leung, H.-T., Shino, S., An, L., Doerge, R.W., Metzstein, M.M., Pak, W.L., et al. (2010). The nonsense-mediated decay pathway maintains synapse architecture and synaptic vesicle cycle efficacy. J. Cell Science.

Longman, D., Plasterk, R.H.A., Johnstone, I.L., and Cáceres, J.F. (2007). Mechanistic insights and identification of two novel factors in the C. elegans NMD pathway. Genes Dev. *21*, 1075–1085.

Metzstein, M.M., and Krasnow, M.A. (2006). Functions of the nonsense-mediated mRNA decay pathway in Drosophila development. PLoS Genet. *2*, e180.

Moore, M.J. (2005). From birth to death: the complex lives of eukaryotic mRNAs. Science *309*, 1514–1518.

Rees, D.C., Johnson, E., and Lewinson, O. (2009). ABC transporters: the power to change. Nat. Rev. Mol. Cell Biol. *10*, 218–227.

Rosains, J., and Mango, S.E. (2012). Genetic characterization of smg-8 mutants reveals no role in C. elegans nonsense mediated decay. PLoS ONE *7*, e49490.

Sakaki, K., Yoshina, S., Shen, X., Han, J., DeSantis, M.R., Xiong, M., Mitani, S., and Kaufman, R.J. (2012). RNA surveillance is required for endoplasmic reticulum homeostasis. Proc. Natl. Acad. Sci. USA *109*, 8079–8084.

Shiga, Y., Tanaka-Matakatsu, M., and Hayashi, S. (1996). A nuclear GFP/ß-galactosidase fusion protein as a marker for morphogenesis in living Drosophila. Develop. Growth Differ. 99–106.

Venken, K.J.T., and Bellen, H.J. (2005). Emerging technologies for gene manipulation in Drosophila melanogaster. Nat. Rev. Genet. *6*, 167–178.

Venken, K.J.T., Carlson, J.W., Schulze, K.L., Pan, H., He, Y., Spokony, R., Wan, K.H., Koriabine, M., de Jong, P.J., White, K.P., et al. (2009). Versatile P[acman] BAC libraries for transgenesis studies in Drosophila melanogaster. Nat. Chem. Biol. *6*, 431–434.

Venken, K.J.T., He, Y., Hoskins, R.A., and Bellen, H.J. (2006). P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. Science *314*, 1747–1751.

Weaver, M., and Krasnow, M.A. (2008). Dual origin of tissue-specific progenitor cells in Drosophila tracheal remodeling. Science *321*, 1496–1499.

Xu, J., Kiel, M.C., Golshani, A., Chosay, J.G., Aoki, H., and Ganoza, M.C. (2006). Molecular localization of a ribosome-dependent ATPase on Escherichia coli ribosomes. Nucl. Acids Res. *34*, 1158–1165.

Xu, T., and Rubin, G.M. (1993). Analysis of genetic mosaics in developing and adult Drosophila tissues. Development *117*, 1223–1237.

Yamashita, A., Izumi, N., Kashima, I., Ohnishi, T., Saari, B., Katsuhata, Y., Muramatsu, R., Morita, T., Iwamatsu, A., Hachiya, T., et al. (2009). SMG-8 and SMG-9, two novel

subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. Genes Dev. *23*, 1091–1105.

CHAPTER 4

IN VIVO ANALYSIS OF *SMG6* DECAY MECHANISMS

Introduction

Smg6 was originally identified along with five other *Smg* genes, in an informational suppressor screen performed in *C. elegans* (Hodgkin et al., 1989), and was subsequently shown to be required for destabilization of nonsense-containing transcripts (Pulak and Anderson, 1993). Orthologs of *Smg6* are conserved in *Drosophila* and mammalian genomes and are known to be required for NMD in both *Drosophila* S2 and mammalian cells (Eberle et al., 2009; Gatfield et al., 2003; Huntzinger et al., 2008). *Smg6*, along with *Smg5* and *Smg7*, two other genes required for NMD, contains a PIN domain (Chiu et al., 2003; Gatfield et al., 2003). The PIN domain is an evolutionarily conserved endonuclease domain that functions in diverse RNA processing enzymes (Schoenberg, 2011). However, both the *Smg5* and *Smg7* PIN domains contain amino acid substitutions in the conserved triad of aspartic acid residues known to be required for catalytic function (Glavan et al., 2006; Takeshita et al., 2007).

Endonucleolytic cleavage of transcripts targeted by nonsense mediated decay has been shown in *Drosophila* and mammalian (Eberle et al., 2009; Huntzinger et al., 2008) cell lines to be carried out by the NMD-specific nuclease Smg6. In *Drosophila* cell culture, cleavage of PTC-containing mRNA is initiated by Smg6-dependent cleavage that

occurs near the premature termination codon (Gatfield and Izaurralde, 2004). This endonucleolytic cleavage is attributed to the catalytic function of Smg6 in both *Drosophila* and mammalian cells (Eberle et al., 2009; Hodgkin et al., 1989; Huntzinger et al., 2008) and has also been shown to occur in zebrafish (Pulak and Anderson, 1993; Wittkopp et al., 2009).

Prior to the discovery that endonucleolytic cleavage was involved in NMD, other degradation steps were thought to be important for NMD. In yeast, which lacks a *Smg6* ortholog, rapid degradation of NMD targets is initiated by deadenylation-independent decapping followed by 5'-3' exonucleolytic degradation by XRN1 (Eberle et al., 2009; Gatfield et al., 2003; He and Jacobson, 2001; Huntzinger et al., 2008). Experimental inactivation of XRN1 revealed a compensatory mechanism dependent on a rapid deadenylation pathway (Chiu et al., 2003; Gatfield et al., 2003; Muhlrad and Parker, 1994). Mammalian cells have also been shown to utilize deadenylation-independent decapping coupled to NMD (Lejeune et al., 2003; Schoenberg, 2011), and a rapid deadenylation pathway (Chen and Shyu, 2003; Couttet and Grange, 2004; Glavan et al., 2006; Lejeune et al., 2003; Takeshita et al., 2007). Neither deadenylation independent decapping or rapid deadenylation have been observed in *Drosophila* S2 cells, and only endonucleolytic cleavage fragments of NMD targeted transcripts have been detected (Eberle et al., 2009; Gatfield and Izaurralde, 2004; Huntzinger et al., 2008). Endonucleolytic cleavage by Smg6 occurs in both *Drosophila* and mammalian cells, however, the lack of NMD-dependent deadenylation-independent decapping and rapid deadenylation activity in *Drosophila* suggests more complex organisms may have co-opted multiple degradation pathways for NMD.

Our data in *Drosophila* suggest degradation of NMD-targeted transcripts may be more complicated *in vivo* than has been observed in *Drosophila* S2 cells, as loss of *Smg6* only causes a modest NMD defect *in vivo* (Frizzell et al., 2012; Gatfield and Izaurralde, 2004). If cleavage by Smg6 were the only nuclease pathway utilized in *Drosophila*, a much more severe NMD defect might be expected as transcripts targeted by NMD would be incapable of the accelerated decay that results from endonucleolytic cleavage. The modest NMD defect associated with *Smg6* mutants suggests other pathways, or partially redundant endonucleases, are utilized for degradation in *Drosophila*. One hypothesis is that Smg6 exonuclease activity is the only cleavage mechanism used in *Drosophila* S2 cells, and that different cell types within *Drosophila* utilize other degradation pathways like those seen in mammalian cells. Another, nonexclusive, hypothesis is that endonucleolytic cleavage by Smg6 is the primary method of degradation, but that other compensatory pathways can initiate cleavage in the absence of Smg6 function. We can use the *Smg6* allelic series identified in our screen (Frizzell et al., 2012) to begin testing these hypotheses *in vivo*.

Additionally, *Smg6* may be an important factor to explain the mechanistic differences of NMD targeting seen in vertebrate lineages. Smg6 is recruited to target transcripts by interactions between its 14-3-3-like domain and phosphorylated Upf1 bound to target mRNA (Okada-Katsuhata et al., 2011). The C-terminal region of vertebrate *Smg6* contains two conserved domains in vertebrates, termed the exon-junction complex binding domains (EBMs) (Kashima et al., 2010). The presence of an exon junction complex downstream of a PTC stimulates NMD in mammalian cells (Buchwald et al., 2010; Hwang et al., 2010; Singh et al., 2008), but does not appear to have any

affect on NMD in *Drosophila* (Gatfield et al., 2003). Indeed, *Drosophila Smg6* lacks EBM domains, suggesting the absence of EJC-dependent NMD stimulation in *Drosophila* is due to a loss, or nonacquisition, of these domains during *Drosophila* evolution.

Here, I present further *in vivo* characterization of the role of *Smg6* in *Drosophila*. First, I investigate the phenotypes associated with *trans*-heterozygous combinations of the *Smg6* alleles. Next, I establish an *in vivo* assay to map the location and efficacy of Smg6 cleavage of NMD targets. Finally, I discuss ways to further analyze *Smg6 in vivo*, in particular, the importance of separating out potentially redundant pathways of target degradation *in vivo.*

Materials and methods

Fly stocks

D. melanogaster lines were raised at 25°C on cornmeal/dextrose medium using standard conditions. Control chromosomes used for the screens were *y w* FRT^{19A} and FRT^{82B} (Xu and Rubin, 1993). *Pacman* RNAi (III) y^1 sc^* v^1; *P{y[+t7.7] v[+t1.8]=TRiP.HMS01169}attP2*, the *daughterless* driver (III) w^{1118}; *P{da-GAL4.w[-]}3/TM3* Sb^1, *P{Act5C-GAL4-w}E1/CyO*, and *Ub-GAL4* were obtained from the Bloomington *Drosophila* stock center. Control crosses utilized *Smg6* alleles from our 3R screen *w;* FRT^{82B} $Smg6^{2a}$*/ TM6B, P{Dfd-EYFP}* Sb^1 ca^1, and included *w;* FRT^{82B} $Smg6^{292}$*/ TM6B, P{Dfd-EYFP}* Sb^1 ca^1 (Frizzell et al., 2012). The hypomorphic allele of the *Drosophila* homolog of XRN1, called *pacman*, pcm^5, was a gift from Sarah Newbury (Grima et al., 2008; Jones et al., 2012; Zabolotskaya et al., 2008). For the balancer

chromosomes we used *w*; ry^{506} Dr^{1}/ TM6B, P{Dfd-EYFP} Sb*^{1} *ca*^{1} (Le et al., 2006). Complementation tests and deficiency experiments for *Smg6* were performed using *Df(3R)ED6220/TM6c Tb, Sb* (Ryder et al., 2007).

Trans-heterozygous combinations

To test the viability of *Smg6* alleles in *trans*, female *FRT*^{82B}, or *FRT*^{82B} *Smg6*/TM6B, P{Dfd-EYFP} Sb*^{1} *ca*^{1} flies were mated to *FRT*^{82B} *Smg6*/TM6B, P{Dfd-EYFP} Sb*^{1} *ca*^{1} males. Progeny adults were analyzed for the presence and absence of the *Sb* marker associated with the balancer chromosome.

pacman knockdown

To map the location of endonucleolytic cleavage of NMD targets by Smg6 in *Drosophila*, the 5'-3' exonuclease, *pacman* (*pcm*), was knocked down by RNAi. Females carrying the *daughterless* driver *P{da-GAL4.w[-]}3* were mated to males carrying the *pcmRNAi* construct *y*^{1} *sc** *v*^{1}*; P{y[+t7.7] v[+t1.8]=TRiP.HMS01169}attP2/TM3 Sb*^{1}; hereafter this RNAi construct will be called *pcmRNAi*. *Sb*$^{+}$ adult male progeny were collected. In addition to this cross, *Ub-GAL4* and *Act5c-GAL4* females were mated to *P{y[+t7.7] v[+t1.8]=TRiP.HMS01169}attP2 /TM3 Sb*^{1}. *Sb*$^{+}$ adult male progeny were collected. Samples were flash frozen in preparation for RNA isolation and qRT-PCR analysis.

RNA isolation and quantitative RT-PCR

RNA isolation and qRT-PCR were performed as described in (Frizzell et al., 2012). Briefly, total RNA was isolated from five to ten flash-frozen samples of adult males using Trizol reagent (Invitrogen) and phase-lock tubes (5-Prime), and the RNeasy mini kit (Qiagen). Genomic contamination was reduced by on-column DNAse treatment (Ambion). RNA concentration was determined using a spectrophotometer and normalized for reverse transcription using the MuMulv reverse transcriptase Retroscript Kit (Ambion). Quantitative RT-PCR was performed on a BioRad iCycler thermocycler using SYBR green supermix (Bio-Rad). All experimental reactions were performed in triplicate with a minimum of two biological replicates. For quantification samples were normalized to *RpL32*.

Results and discussion

Viability of *trans*-heterozygous *Smg6* alleles

The viability of *Smg6* alleles was measured in *trans* to a deficiency uncovering *Smg6* (Frizzell et al., 2012). As a control, two null alleles, $Smg6^{52a}$ and $Smg6^{292}$, were also mated to verify the combination of two null *Smg6* alleles was not more severe than either allele in *trans* to a deficiency. All animals were collected as *trans*-heterozygous adults. We found the null alleles were about 60% viable, both in *trans* to each other, or in *trans* to a *Smg6* deficiency (Fig. 2.1). We found the viability associated with one of the hypomorphic alleles, $Smg6^{21a}$, did not fit the trend associated with other alleles. While other hypomorphic alleles were 80% viable in *trans* to a deficiency, $Smg6^{21a}$ was only ~60% viable (Fig. 4.1) (Frizzell et al., 2012). To test if the decreased fitness of

Smg6^{21a} was due to the particular mutation in *Smg6* or to another mutation on the *Smg6*^{21a} chromosome, *trans*-heterozygous combinations of the three hypomorphic alleles, *Smg6*^{21a}, *Smg6*^{91a}, and *Smg6*^{2a}, were measured for viability. Interestingly, when the hypomorph, *Smg6*^{21a}, was crossed to either *Smg6*^{91a} or *Smg6*^{2a}, it decreased the viability of the alleles from 80% viable to 60% viable. However, *Smg6*^{91a} in *trans* to *Smg6*^{2a} was nearly 90% viable. This result suggests *Smg6*^{21a} acts as a null allele when scored for viability. The lesion in *Smg6*^{21a} is a premature termination codon only a few base pairs downstream of the lesion in *Smg6*^{91a}, and previous work showed this allele has a hypomorphic NMD defect (Frizzell et al., 2012). It is more likely that the *Smg6*^{21a} chromosome has picked up a second mutation that is haploinsufficient for viability, but does not affect NMD. In order to more accurately judge whether this defect in viability is due to the *Smg6* mutation or to a second site effect, more *trans*-heterozygous combinations need to be made; in particular, *trans*-heterozygous null alleles over hypomorphic alleles need to be tested. If *Smg6*^{21a} has an extra mutation on the chromosome that affects viability, a greater decrease in viability in the *Smg6*^{21a} /null combination would be expected.

To further test the roles of the partial and full loss of *Smg6* function in these alleles, more direct tests of NMD function must be done, including testing the stability of known NMD targets in these *trans*-heterozygous combinations.

Mapping Smg6 cleavage using *pacman* knockdown

To test the site specificity and efficacy of Smg6 cleavage of NMD targeted transcripts, we modified an assay that has previously been performed in tissue culture (Franks et al., 2010; Gatfield and Izaurralde, 2004; Rehwinkel et al., 2005) to detect the

cleavage products that result from Smg6-mediated cleavage of an NMD target. Smg6 endonucleolytic activity has been demonstrated in *Drosophila* S2 (Huntzinger et al., 2008) and mammalian cell culture (Eberle et al., 2009) to cleave NMD-targeted transcripts near the premature termination codon (Gatfield and Izaurralde, 2004). This cleavage frees the 3' end of the 5' fragment for degradation by the exosome, and the 5' end of the 3' fragment for degradation by the exonuclease, *pacman* (*pcm*), called XRN1 in other organisms (Gatfield and Izaurralde, 2004; He and Jacobson, 2001). We set out to determine if, following endonucleolytic cleavage, *pcm* and exosome mediated degradation occurs in intact *Drosophila*. We did this by measuring relative expression levels of the 5' and 3' degradation fragments of targets in *Smg6* mutant backgrounds, where both fragments should be stabilized. We also measured relative expression in *pcm* knockdown animals, where the 3' fragment should be stabilized relative to the 5' fragment, whose 3' end is still susceptible to degradation by the exosome.

This assay can also be used to address whether cleavage occurs on endogenous transcripts that harbor no premature termination codon, but are nonetheless targeted for degradation by NMD. To detect where the cleavage occurs, and whether the cleavage is dependent on Smg6 function, we first designed primers to well-characterized NMD targets. The primers were designed to detect amplicons that lie 5' or 3' to the predicted cleavage site of either a premature or a native termination codon (Fig. 4.2 B). For this experiment two different NMD targets were utilized. The first target, *transformer* (*tra*), is a sex-specific splicing regulator that is alternatively spliced into a PTC-containing isoform in males, and represents our PTC-containing transcript (Metzstein and Krasnow, 2006). Stabilization of *tra* fragments has been previously seen in *Drosophila* S2 cells

using northern analysis (Rehwinkel et al., 2005). The second target is *Gadd45*. *Gadd45* contains no introns or splice forms that would make it an obvious NMD target; however, it has a relatively long 3' UTR and is a direct NMD target in *Drosophila* (A. Chapin, in prep) and mammals (Viegas et al., 2007). *Gadd45* represents our endogenous NMD target that lacks a PTC. Primers for the *Gadd45* transcript were designed 5' and 3' to the native termination codon. Both *tra* and *Gadd45* were stabilized relative to wild-type controls in microarrays of $Smg6^{292}$ and $Smg6^{2a}$, and were stabilized in microarrays of other NMD genes, $Upf2^{25G}$ and $Smg1^{32AP}$ (Frizzell et al., 2012; Metzstein and Krasnow, 2006). Alleles of *Upf2* and *Smg1* are used as controls for NMD function in the following experiments.

To determine relative levels of *tra* 5' and 3' amplicons, we first determined if both amplicons were stabilized in *Smg6* mutants. We predicted that the endonucleolytic cleavage does not occur in *Smg6* mutants, therefore the relative levels of the amplicons should be stabilized to approximately equal levels as compared to a wild-type control. However, this analysis could be confounded if another degradation pathway is utilized. We found that in $Smg6^{292}$, a null *Smg6* allele, both the *tra* 5' and *tra* 3' amplicons were stabilized about 1.5 fold over the same amplicons in a wild-type background (Fig. 4.2 C, D). Interestingly, in $Smg6^{2a}$, a hypomorphic allele that harbors a missense mutation in the endonuclease domain of *Smg6* (Frizzell et al., 2012), the 3' amplicon was stabilized more, about 1.8 fold relative to the 5' amplicon, which was barely stabilized over wild-type controls (Fig. 4.2 C, D). This indicates that the hypomorphic mutation in $Smg6^{2a}$ somehow stabilizes the 3' fragment, while the 5' end is degraded as if NMD was not disrupted. As a control we also analyzed the levels of the 5' and 3' *tra* amplicons in male

animals hemizygous for *Smg1*^{32AP}, which has a weak NMD defect, and *Upf2*^{25G}, which has a complete loss of NMD function. Both 5' and 3' amplicons were stabilized about 1.5-2.5 fold in *Smg1*^{32AP}, and there was no significant difference in the overall expression level between the amplicons (Fig. 4.2 C, D). In *Upf2*^{25G}, there was a significant stabilization of the 3' amplicon relative to the 5' amplicon (Fig. 4.2 C, D).

To determine if endonucleolytic cleavage of *tra* occurs at the isoform-specific premature termination codon we assayed whether the 3' amplicon was stabilized relative to the 5' amplicon in male animals ubiquitously expressing *pcmRNAi* from the *da-GAL4* promoter. Two other drivers were tested for this experiment. The first, *Act5c-GAL4* did not result in a detectable knockdown of *pcm* transcript (data not shown). A second driver, *Ub-GAL4,* resulted in adult lethality, a result that is not surprising as null alleles of *pacman* result in pupal lethality (Jones et al., 2012). Male flies expressing *da-GAL4/UAS:pcmRNAi* displayed a nearly two-fold reduction in *pacman* transcript (Fig. 4.2 A). In these animals, the 3' *tra* fragment was stabilized two-fold relative to the 5' *tra* fragment (Fig. 4.2 C, D). This indicates the cleavage does occur somewhere between the two amplicons and that this method can detect 3' fragments stabilized upon loss of the 5' exonuclease, pacman.

Next, we determined if *Gadd45* is cleaved by Smg6 in the vicinity of its normal termination codon. Using the same approach as the *tra* experiment, the relative levels of *Gadd45* 3' and *Gadd45* 5' fragments were measured in wild type animals along with *Smg6*^{2a}, and *Smg6*^{292} mutants. Interestingly, in both the *Smg6* mutants, the 5' fragment was stabilized two-fold relative to the 3' fragment (Fig.4.2 E, F). The stabilization of the 5' fragment relative to the 3' fragment was also seen in *Upf2*^{25G} mutants, but was not

observed in *Smg1*^{32AP} mutants where they were not stabilized at all (Fig. 4.2 E, F). The difference in stabilization of the *Gadd45* amplicons could be due to many reasons. First, the efficiency of primer binding to the transcripts could introduce technical errors with detection of the transcripts; more primer pairs to the different amplicons should be used to recapitulate these results. Second, the 5' amplicon stabilization is strange in light of the 3' stabilization seen in the *tra* amplicons. Perhaps this difference is a result of a premature versus native termination codon targeted by NMD. As predicted by our model, when *pacman* is knocked down using RNAi, the 3' fragment is stabilized relative to the 5' fragment (Fig. 4.2 E, F). This indicates that *Gadd45* is cleaved near the normal termination codon. However, the degree of *Smg6* cleavage necessary for this stabilization will need to be determined using double mutant analysis.

To ensure the phenotype is due to Smg6-mediated endonucleolytic cleavage of *tra* or *Gadd45* the 5' and 3' fragments should be assayed in a *Smg6* and *pacman* double mutant. Unfortunately both the *da-GAL4* driver and *pcmRNAi* are inserted on the third chromosome, which makes an RNAi, mutant/deficiency double mutant technically challenging. However, this experiment can be done in one of two ways; first, the *da-GAL4* promoter can be used to knock down both *pcm* and *Smg6*. The *Smg6RNAi* has been tested for an NMD defect and it causes stability of an NMD reporter (A. Chapin personal communication), however, the 5' and 3' amplicons would have to be characterized in this background before an experiment can be performed. Second, *pcm5; Smg6*/Df* mutant can be made, however, the *pcm5* allele has not yet been tested using these methods and needs to be verified. Additionally, the *pcm5* allele is temperature sensitive (Grima et al., 2008) and may be sick when combined with *Smg6*. To perform

this experiment efficiently, male L3 larvae of the double mutant should be collected to avoid developmental defects associated with loss of Smg6 and pcm. However, all experiments would need to be optimized in L3 larvae.

This analysis of *Smg6* provides the basis for more assays of NMD function *in vivo* in *Drosophila*. First, measuring viability in *trans*-heterozygous combinations of hypomorphs indicates that additional factors on one of the chromosomes may dominantly affect viability without having a significant affect on NMD function (Frizzell et al., 2012). More *trans*-heterozygous combinations and determining the NMD defects in these combinations can provide more information on the roles of different *Smg6* domains on NMD function.

The XRN1-based assay developed in this chapter has been used extensively in tissue culture (Behm-Ansmant et al., 2007; Eberle et al., 2009; Franks et al., 2010; Gatfield and Izaurralde, 2004; Huntzinger et al., 2008; Lejeune et al., 2003), with the assumption that it provides a read-out of direct versus indirect NMD targets (Rehwinkel et al., 2005). Upon loss of XRN1, direct targets are predicted to display stabilization of the 3' fragment; however, general loss of NMD function should result in direct targets showing equal fragment amplification. In contrast, indirect targets will not show amplicon specific stabilization. Applying this assay to mutants, instead of RNAi knockdown of NMD components, adds another level of complexity to the experiment. Hypomorphic mutants such as $Smg6^{2a}$ and $Upf2^{25G}$ may have residual function that allows the NMD complex to assemble on targeted mRNA, but may prevent dissociation of the NMD mRNP leading to stabilization of the fragment that is bound by the NMD complex. Indeed, this type of stabilization has been observed in mammalian cells upon

loss of Upf1 hydrolysis of ATP (Franks et al., 2010). In these experiments, expressing Upf1 mutants that were unable to hydrolyze ATP resulted in components of the NMD complex remaining bound to each other and to the 3' fragment of PTC-containing mRNA, this resulted in protection of the 3' fragment from degradation by the 5' exonuclease, XRN1 (Franks et al., 2010). One possibility is that *Upf2*^{25G} is behaving similarly to the Upf1 mutants. *Upf2*^{25G} mutants have a strong NMD defect (Avery et al., 2011; Frizzell et al., 2012; Metzstein and Krasnow, 2006) and the lesion in *Upf2*^{25G} changes the natural stop codon to an arginine, extending the Upf2 protein by fifteen residues (Metzstein and Krasnow, 2006). The extra C-terminal amino acids added to the *Upf2*^{25G} polypeptide could hinder interactions between *Upf1* and *Upf2*. These are required to stimulate ATP hydrolysis by Upf1 (Clerici et al., 2009) which is required for mRNP remodeling and recycling of NMD components (Franks et al., 2010). In the presence of *Upf2*^{25G} protein, the mRNP might remain associated with the end of the transcript, protecting that fragment from degradation, as has been observed in mammalian cells (Franks et al., 2010). The mRNP association may even vary between transcripts based on the efficiency or mechanism of recognition. This observation might explain why *tra* transcripts, which encode a PTC, have a stabilized 3' fragment in *Upf2*^{25G} mutants while *Gadd45* transcripts, which do not have an obvious PTC, have the 5' fragments stabilized in *Upf2*^{25G} mutants.

The idea that the NMD mRNP could protect cleaved transcripts is also interesting in light of our data that shows a *Smg6* hypomorph predicted to contain a catalytically dead endonuclease (Glavan et al., 2006; Takeshita et al., 2007) domain may have residual endonuclease activity. In the null allele, *Smg6*^{292}, amplicons representing the 5' and 3' cleavage fragments are stabilized equally, as would be predicted in the complete absence

of endonuclease function. However, in the hypomorph *Smg6*^{2a}, we found that the 3' amplicon was stabilized relative to the 5' amplicon (Fig. 2 C, D). The lesion in *Smg6*^{2a} converts a conserved catalytic residue in the PIN endonucleolytic domain from an aspartic acid to a valine, presumably disrupting the function of the catalytic domain, but likely leaving the overall structure of the protein intact. The stabilization of the 3' amplicon suggests the mutation may allow Smg6^{2a} to assemble with the NMD complex and that inefficient cleavage, due to the active site mutation, might result in inefficient mRNP remodeling and thus, protection of the 3' *tra* fragment. However, these results are not mirrored in the *Gadd45* amplicons, where the 5' amplicon is stabilized relative to the 3' amplicon in both *Smg6*^{2a} and *Smg6*^{292} mutants (Fig. 4.3 E, F). This is intriguing because the *Gadd45* transcript may be targeted differently from *tra* because it lacks a PTC. These observations generate models that can be tested using the aforementioned cleavage assays.

Future directions

Work on this cleavage assay for determining relative levels of cleavage products of NMD targets is still in a preliminary phase of optimization. However, several other experiments using this assay can be hypothesized. Additionally, alleles of NMD genes can be utilized for further characterization of the role of NMD *in vivo*. This section describes some such future experiments.

First, to confirm results and to examine if there is a trend in stabilization of cleavage products in hypomorphic NMD mutants, all the *pcmRNAi* experiments need to validated using additional primer pairs against the 5' and 3' cleavage products of *tra*,

Gadd45, and other targets. Additional targets that may be informative include *Smg5*, a known NMD component and NMD target (Rehwinkel et al., 2005; Yepiskoposyan et al., 2011), and fluorescent reporters containing the SV40 3'UTR. Both *Smg5* and the SV40 3'UTR are targets that do not contain a PTC (Metzstein and Krasnow, 2006). *Bona fide* PTC-containing alleles should also be assayed, such as Adh^{n4} (Chen et al., 2005; Chia et al., 1987; Metzstein and Krasnow, 2006), to determine if there is a difference in stabilization between different types of NMD targets, such as PTC-containing versus endogenous targets.

Further characterization of cleavage products can be performed in a *pacman* mutant, as opposed to RNAi, as was described in this chapter. We have obtained a hypomorphic temperature sensitive allele of pacman, pcm^5 (Zabolotskaya et al., 2008). This allele is viable at permissive temperatures, so characterization of cleavage fragments at any developmental stage should be possible (Grima et al., 2008). One benefit of using a mutant, as opposed to RNAi, is that all cells are defective, which allows for analysis of the effect of loss of function of *pcm* in whole animals. Also, the pcm^5 mutation is present on the X chromosome, so double mutant analysis with *Smg6* or other NMD mutants can be performed using hemizygous pcm^5 males. However, characterization of the pcm^5 still needs to be performed in our laboratory.

In addition to using an allele of *pcm*, we should look at NMD cleavage products in different NMD mutant backgrounds. Alleles of *Smg5*, for example, may be interesting as *Smg5*-dependent dephosphorylation has been shown biochemically to promote remodeling of the NMD mRNP (Chiu et al., 2003; Okada-Katsuhata et al., 2011).

The stabilization of either the 3' fragment of *tra* or the 5' fragment of *Gadd45* in *Smg6* mutants is intriguing and could be followed up on using protein-based assays. One experimental option is to use either UV or formaldehyde crosslinking followed by immunoprecipitation of NMD proteins and qRT-PCR to identify sequences associated with the NMD mRNP (Darnell, 2012). If this can be used quantitatively, we can test if the mRNP is associating upstream or downstream of the predicted cleavage position.

Since we can use probes, such as qRT-PCR or northern blot analysis, to identify differences in stability of an NMD target that has undergone endonucleolytic cleavage by Smg6, we may be able to expand our assays to identify more direct targets of NMD. Detection of endonucleolytic cleavage products could easily be expanded to either RNA-sequencing, or to transcripts that have multiple probes on a microarray that lie throughout the transcript. We already have microarray data for the two alleles of *Smg6* that were used in this experiment (Frizzell et al., 2012); further analysis could include reannotating the array to identify probes that could detect the 5' and 3' ends of transcripts. Once this is done, microarrays on *pacman* RNAi or mutants and *pacman*; *Smg6* double mutants could be an efficient way to detect transcripts directly targeted by Smg6. These could then be verified by qRT-PCR.

Another aspect of the NMD screen is that, while it efficiently identifies genes with a primary role in NMD, it might miss genes that provide a secondary or compensatory role. Of particular note, *Smg6* mutations result in a minor NMD defect (Frizzell et al., 2012), suggesting other degradation pathways are utilized by NMD. In yeast (He and Jacobson, 1995) and mammalian cells (Chen and Shyu, 2003; Couttet and Grange, 2004; Muhlrad and Parker, 1994), deadenylation independent decapping and

rapid deadenylation contribute to the rapid decay characteristic of NMD. However, this has not been demonstrated to occur in *Drosophila* S2 cells (Gatfield and Izaurralde, 2004). Potentially, deadenylation and decapping, or novel ribonuclease pathways, may be used as a secondary degradation mechanism *in vivo*. Deadenylation and decapping coupled to NMD may not be active in NMD in S2 cells, or could be cell, tissue, or developmental stage dependent and so would have been missed when analyzed in tissue culture. We could perform a screen looking for secondary decay pathways that are utilized by NMD. One way to do this would be to use a candidate-based RNAi screen. There are 200 genes in *Drosophila* that contain the GO-term "ribonuclease". Forty-three of these ribonuclease genes have identified alleles or previously made RNAi constructs directed against them. A whole organism and tissue directed screen looking for synthetic lethality between alleles of *Smg6* and RNAi against other candidate ribonucleases could be performed using such reagents. A ubiquitous driver, as well as a nervous system specific driver, that are sufficient for rescue of NMD-based lethality of *Upf2* (A. Chapin, personal communication), could be used to drive RNAi in a viable *Smg6* background; RNAi lines that result in synthetic lethality using both drivers could then be tested for NMD reporter enhancement. This method may uncover genes that would have been missed due to *Smg6* being the primary nuclease used by NMD, therefore a functioning *Smg6* would have compensated for minor NMD defects resulting from loss of other ribonucleases.

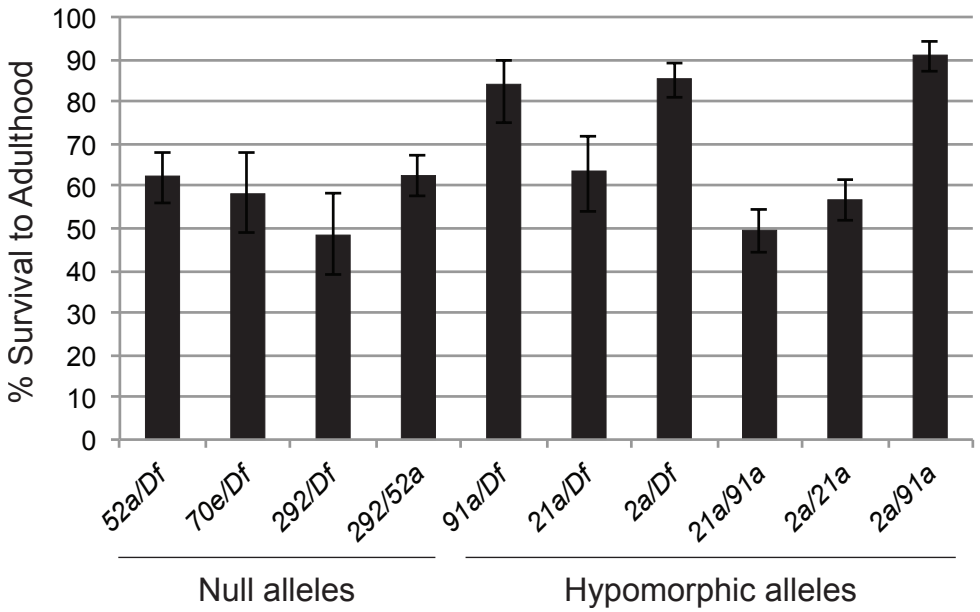
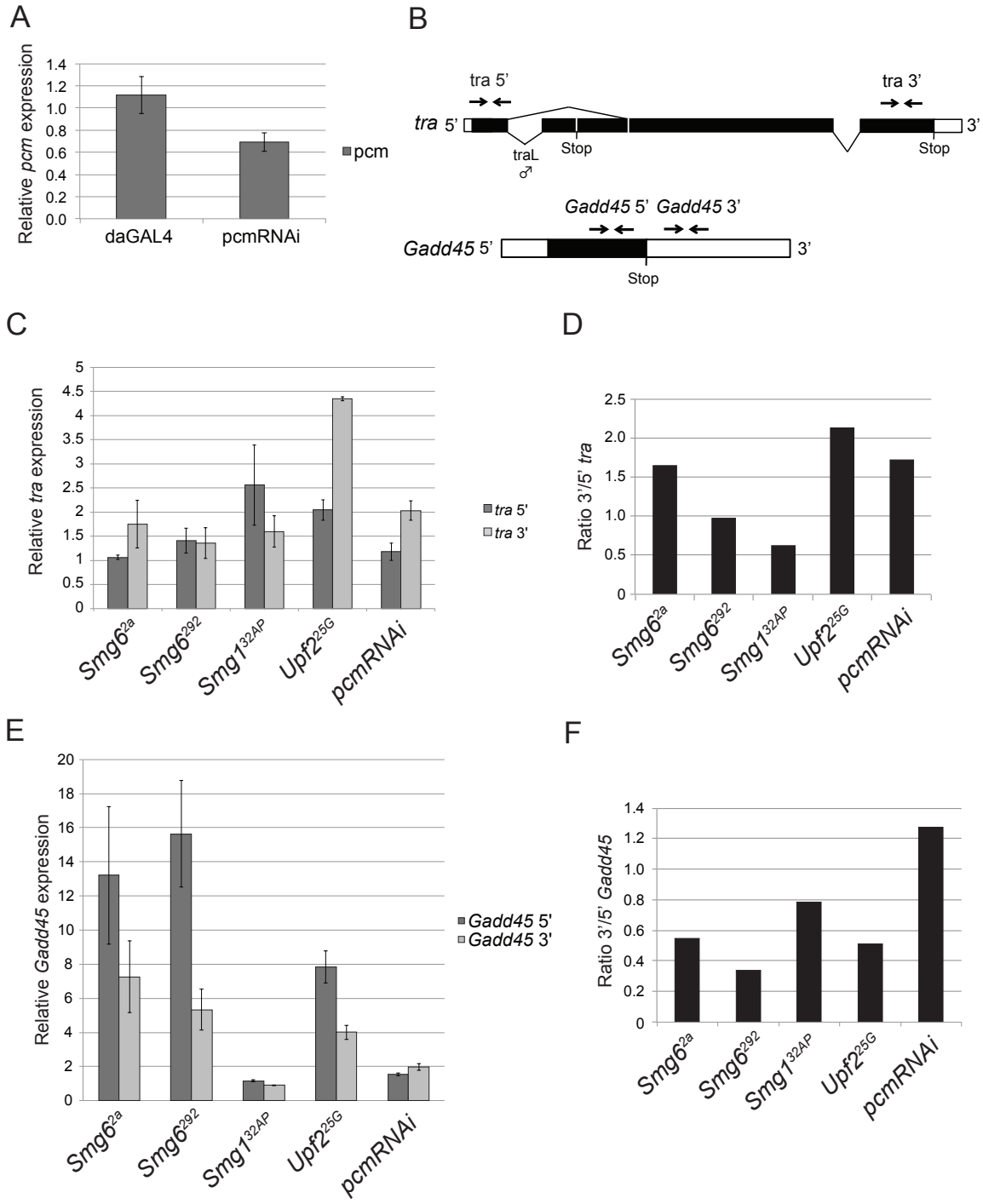


Figure 4.1 Viability of *Smg6* mutants. Percentage survival to adulthood of *Smg6 trans*-heterozygous combinations. Alleles shown in *trans* to a deficiency are percentages first presented in (Frizzell et al., 2012). Alleles are classified as null or hypomorphic based on molecular criteria. Percentage survival is calculated relative to balancer containing siblings. Error bars: 95% confidence interval of the binomial distribution.

Figure 4.2 Analysis of *Smg6* cleavage products and *pacman* knockdown. (A) Levels of *pcm* transcript in *daGAL4* control animals and *daGAL4*/*pcmRNAi* animals. Error bars indicate 1X S.E.M, n=4. (B) NMD target transcripts and location of 5' and 3' primers in *tra* and *Gadd45*. (C) Effect of *Smg6*, *Smg1, Upf2* mutations and knockdown of *pcm* on levels of the 5' and 3' cleavage fragments of the endogenous NMD target *tra* in males. Levels for *Smg6* alleles are normalized to FRT^{82B}/Df animals; $Smg1^{32AP}/Y$ and $Upf2^{25G}/Y$ are normalized to FRT^{19A}/Y. *daGAL4*/*pcmRNAi* animals are normalized to *daGAL4*. Error bars indicate S.E.M., n=2 for *Smg6*, *Upf2*, and *Smg1;* n=4 for *daGAL4* and *daGAL4*/*pcmRNAi*. (D) Fold change of *tra* 3' over *tra* 5' expression in *Smg6*, *Smg1*, *Upf2*, and *pcmRNAi* animals. (E) Effect of *Smg6*, *Smg1, Upf2* mutations and knockdown of *Gadd45* on levels of the 5' and 3' cleavage fragments of the endogenous NMD target *Gadd45* in males. Levels for *Smg6* alleles are normalized to FRT^{82B}/Df animals; $Smg1^{32AP}/Y$ and $Upf2^{25G}/Y$ are normalized to FRT^{19A}/Y. *daGAL4*/*pcmRNAi* animals are normalized to *daGAL4*. Error bars indicate S.E.M., n=2 for *Smg6*, *Upf2*, and *Smg1;* n=4 for *daGAL4* and *daGAL4*/*pcmRNAi*. (D) Fold change of *Gadd45* 3' over *Gadd45* 5' expression in *Smg6*, *Smg1*, *Upf2*, and *pcmRNAi* animals.

Figure 4.2 Continued.



References

Avery, P., Vicente-Crespo, M., Francis, D., Nashchekina, O., Alonso, C.R., and Palacios, I.M. (2011). Drosophila Upf1 and Upf2 loss of function inhibits cell growth and causes animal death in a Upf3-independent manner. Rna *17*, 624–638.

Behm-Ansmant, I., Gatfield, D., Rehwinkel, J., Hilgers, V., and Izaurralde, E. (2007). A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. Embo J. *26*, 1591–1601.

Buchwald, G., Ebert, J., Basquin, C., Sauliere, J., Jayachandran, U., Bono, F., Le Hir, H., and Conti, E. (2010). Insights into the recruitment of the NMD machinery from the crystal structure of a core EJC-UPF3b complex. Proc. Natl. Acad. Sci. USA *107*, 10050–10055.

Chen, C.-Y.A., and Shyu, A.-B. (2003). Rapid deadenylation triggered by a nonsense codon precedes decay of the RNA body in a mammalian cytoplasmic nonsense-mediated decay pathway. Mol. Cell. Biol. *23*, 4805–4813.

Chen, Z., Smith, K.R., Batterham, P., and Robin, C. (2005). Smg1 nonsense mutations do not abolish nonsense-mediated mRNA decay in Drosophila melanogaster. Genetics *171*, 403–406.

Chia, W., Savakis, C., Karp, R., and Ashburner, M. (1987). Adhn4 of Drosophila melanogaster is a nonsense mutation. Nucl. Acids Res. *15*, 3931.

Chiu, S.-Y., Serin, G., Ohara, O., and Maquat, L.E. (2003). Characterization of human Smg5/7a: a protein with similarities to Caenorhabditis elegans SMG5 and SMG7 that functions in the dephosphorylation of Upf1. Rna *9*, 77–87.

Clerici, M., Mourão, A., Gutsche, I., Gehring, N.H., Hentze, M.W., Kulozik, A.E., Kadlec, J., Sattler, M., and Cusack, S. (2009). Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2. Embo J. *28*, 2293–2306.

Couttet, P., and Grange, T. (2004). Premature termination codons enhance mRNA decapping in human cells. Nucl. Acids Res. *32*, 488–494.

Darnell, R. (2012). CLIP (cross-linking and immunoprecipitation) identification of RNAs bound by a specific protein. Cold Spring Harb. Protoc. *2012*.

Eberle, A.B., Lykke-Andersen, S., Mühlemann, O., and Jensen, T.H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat. Struct. Mol. Biol. *16*, 49–55.

Franks, T.M., Singh, G., and Lykke-Andersen, J. (2010). Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense-mediated mRNA decay. Cell *143*, 938–950.

Frizzell, K.A., Rynearson, S.G., and Metzstein, M.M. (2012). Drosophila mutants show NMD pathway activity is reduced, but not eliminated, in the absence of Smg6. Rna *18*, 1475–1486.

Gatfield, D., and Izaurralde, E. (2004). Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. Nature *429*, 575–578.

Gatfield, D., Unterholzner, L., Ciccarelli, F.D., Bork, P., and Izaurralde, E. (2003). Nonsense-mediated mRNA decay in Drosophila: at the intersection of the yeast and mammalian pathways. Embo J. *22*, 3960–3970.

Glavan, F., Behm-Ansmant, I., Izaurralde, E., and Conti, E. (2006). Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. Embo J. *25*, 5117–5125.

Grima, D.P., Sullivan, M., Zabolotskaya, M.V., Browne, C., Seago, J., Wan, K.C., Okada, Y., and Newbury, S.F. (2008). The 5"-3" exoribonuclease pacman is required for epithelial sheet sealing in Drosophila and genetically interacts with the phosphatase puckered. Biol. Cell *100*, 687–701.

He, F., and Jacobson, A. (1995). Identification of a novel component of the nonsense-mediated mRNA decay pathway by use of an interacting protein screen. Genes Dev. *9*, 437–454.

He, F., and Jacobson, A. (2001). Upf1p, Nmd2p, and Upf3p regulate the decapping and exonucleolytic degradation of both nonsense-containing mRNAs and wild-type mRNAs. Mol. Cell. Biol. *21*, 1515–1530.

Hodgkin, J., Papp, A., Pulak, R., Ambros, V., and Anderson, P. (1989). A new kind of informational suppression in the nematode Caenorhabditis elegans. Genetics *123*, 301–313.

Huntzinger, E., Kashima, I., Fauser, M., Saulière, J., and Izaurralde, E. (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. Rna *14*, 2609–2617.

Hwang, J., Sato, H., Tang, Y., Matsuda, D., and Maquat, L.E. (2010). UPF1 association with the cap-binding protein, CBP80, promotes nonsense-mediated mRNA decay at two distinct steps. Mol. Cell *39*, 396–409.

Jones, C.I., Zabolotskaya, M.V., and Newbury, S.F. (2012). The 5" → 3" exoribonuclease XRN1/Pacman and its functions in cellular processes and development. Wiley Interdiscip. Rev. RNA *3*, 455–468.

Kashima, I., Jonas, S., Jayachandran, U., Buchwald, G., Conti, E., Lupas, A.N., and Izaurralde, E. (2010). SMG6 interacts with the exon junction complex via two conserved EJC-binding motifs (EBMs) required for nonsense-mediated mRNA decay. Genes Dev. *24*, 2440–2450.

Le, T., Liang, Z., Patel, H., Yu, M.H., Sivasubramaniam, G., Slovitt, M., Tanentzapf, G., Mohanty, N., Paul, S.M., Wu, V.M., et al. (2006). A new family of Drosophila balancer chromosomes with a w- dfd-GMR yellow fluorescent protein marker. Genetics *174*, 2255–2257.

Lejeune, F., Li, X., and Maquat, L.E. (2003). Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. Mol. Cell *12*, 675–687.

Metzstein, M.M., and Krasnow, M.A. (2006). Functions of the nonsense-mediated mRNA decay pathway in Drosophila development. PLoS Genet. *2*, e180.

Muhlrad, D., and Parker, R. (1994). Premature translational termination triggers mRNA decapping. Nature *370*, 578–581.

Okada-Katsuhata, Y., Yamashita, A., Kutsuzawa, K., Izumi, N., Hirahara, F., and Ohno, S. (2011). N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. Nucl. Acids Res. 1–16.

Pulak, R., and Anderson, P. (1993). mRNA surveillance by the Caenorhabditis elegans smg genes. Genes Dev. *7*, 1885–1897.

Rehwinkel, J., Letunic, I., Raes, J., Bork, P., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. Rna *11*, 1530–1544.

Ryder, E., Ashburner, M., Bautista-Llacer, R., Drummond, J., Webster, J., Johnson, G., Morley, T., Chan, Y.S., Blows, F., Coulson, D., et al. (2007). The DrosDel deletion collection: a Drosophila genomewide chromosomal deficiency resource. Genetics *177*, 615–629.

Schoenberg, D.R. (2011). Mechanisms of endonuclease-mediated mRNA decay. Wiley Interdiscip. Rev. RNA *2*, 582–600.

Singh, G., Rebbapragada, I., and Lykke-Andersen, J. (2008). A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. PLoS Biol. *6*, e111.

Takeshita, D., Zenno, S., Lee, W.C., Saigo, K., and Tanokura, M. (2007). Crystal structure of the PIN domain of human telomerase-associated protein EST1A. Proteins *68*, 980–989.

Viegas, M.H., Gehring, N.H., Breit, S., Hentze, M.W., and Kulozik, A.E. (2007). The abundance of RNPS1, a protein component of the exon junction complex, can determine the variability in efficiency of the nonsense mediated decay pathway. Nucl. Acids Res. *35*, 4542–4551.

Wittkopp, N., Huntzinger, E., Weiler, C., Sauliere, J., Schmidt, S., Sonawane, M., and

Izaurralde, E. (2009). Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic development and survival. Mol. Cell. Biol. *29*, 3517–3528.

Xu, T., and Rubin, G.M. (1993). Analysis of genetic mosaics in developing and adult Drosophila tissues. Development *117*, 1223–1237.

Yepiskoposyan, H., Aeschimann, F., Nilsson, D., Okoniewski, M., and Mühlemann, O. (2011). Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. Rna *17*, 2108–2118.

Zabolotskaya, M.V., Grima, D.P., Lin, M.-D., Chou, T.-B., and Newbury, S.F. (2008). The 5“-3” exoribonuclease Pacman is required for normal male fertility and is dynamically localized in cytoplasmic particles in Drosophila testis cells. Biochem. J. *416*, 327–335.

CHAPTER 5

CHARACTERIZATION OF *CIS*-ACTING SEQUENCES IN THE SV40 3'UTR REQUIRED FOR TARGETING BY NMD

Introduction

While performing a forward genetic screen in *Drosophila* for genes required for cell morphogenesis, a serendipitous discovery was made: the fluorescent reporter used to identify homozygous mutant cells was subject to regulation by nonsense mediated decay (NMD) (Metzstein and Krasnow, 2006). NMD pathway mutants show a dramatic up-regulation of the fluorescent reporter expression attributable to stabilization of the reporter transcript. The fluorescent reporter does not contain a premature termination codon (PTC). However, independent features of the reporter were tested for NMD-dependent destabilization and it was determined that the 3'UTR utilized in the fluorescent reporter was specifically targeted by NMD (Metzstein and Krasnow, 2006). This 3'UTR is found in cloning plasmids used in early *Drosophila* molecular genetic experiments (Thummel et al., 1988), and is derived from the *Simian virus 40 (SV40)*, a double stranded DNA virus that infects both primates and humans. The SV40 3'UTR is derived from the small t antigen intron sequence and it is unclear why the SV40 3'UTR is specifically targeted by NMD, but there are a few potential mechanisms. For example, long 3'UTRs have been shown in other systems to trigger NMD (Amrani et al., 2004;

Eberle et al., 2008; Hogg and Goff, 2010; Kertész et al., 2006; Muhlrad and Parker, 1999; Singh et al., 2008). The SV40 3'UTR is 720 bp long and the average length of a *Drosophila* 3'UTR is ~400 bp (Celniker and Rubin, 2003); thus, the length of the SV40 3'UTR may cause it to trigger NMD in a length-dependent manner. However, many endogenous 3'UTRs in *Drosophila* are 400 bp or longer and are not NMD targets (Rehwinkel et al., 2005). This suggests that targeting of 3'UTRs, and in particular, this reporter, is more complicated than a pure length-measuring mechanism.

Another potential mechanism depends on the exon junction complex (EJC) promoting NMD, as is seen in plants (Kerényi et al., 2008; Kertész et al., 2006), zebrafish (Wittkopp et al., 2009), and mammalian cells (Buchwald et al., 2010; Carter et al., 1996; Le Hir et al., 2001; Maquat, 2004; Zhang et al., 1998). The SV40 3'UTR contains an inefficiently spliced intron (M. Metzstein personal communication), which could stimulate NMD in an EJC-dependent manner. However, the EJC does not on its own stimulate NMD targeting in *Drosophila* S2 cells (Gatfield et al., 2003), and loss of EJC components does not result in a visible increase in SV40 3'UTR-based reporter enhancement (J. Nelson personal communication). This suggests EJC-dependent targeting is not the primary mechanism of targeting the SV40 3'UTR. Additionally, deleting the intron sequence did not result in a significant reduction in NMD targeting, based on a visual assay of reporter expression levels (Metzstein and Krasnow, 2006). However, EJC-dependent NMD may represent only one of multiple NMD pathways and therefore be dependent on the individual NMD mutant used for testing the reporter expression. Additionally, visual fluorescence may not be reliably quantitative, whereas qRT-PCR provides a more quantitative assay of transcript stability. Another possibility

is that the intron only plays a minor stimulatory role in *Drosophila*; thus, detection of this stimulation may be difficult to resolve if length-dependent targeting is intact.

Finally, the SV40 3'UTR may be targeted by NMD due to specific sequence features. The SV40 3'UTR is derived from a virus. Many viral messages are sensitive to NMD because they contain long, polycistronic messages that are inefficiently spliced (Withers and Beemon, 2010a). Polycistronic messages appear similar to PTC-containing transcripts since the multiple open reading frames look like upstream open reading frames, the first termination codon of which is identified as a PTC, making the rest of the transcript appear as a long 3'UTR. Unspliced or inefficiently spliced transcripts may be marked with exon junction complexes downstream of the polycistronic termination codons, and since the EJC is thought to stimulate NMD targeting (Maquat, 2004) the viral transcript should be rapidly degraded by NMD. Interestingly, at least one example of a viral transcript containing these NMD-target features is known that actually evades NMD. The Rous sarcoma virus (RSV) is an avian retrovirus that, when transcribed by the host cellular machinery, encodes a polycistronic transcript with inefficiently spliced introns. The polycistronic transcript mimics a long 3'UTR and the introns may have EJC components associated after intron splicing, two features that make this viral transcript a potential NMD target. However, RSV evades NMD through a mechanism that is dependent on a ~400 nt structured RNA segment called the Rous stability element (RSE) (Withers and Beemon, 2010b; 2011). While the mechanism by which the RSE evades NMD is unclear, its structure might allow for folding of the capped and polyadenylated transcript such that the poly-A-tail lies in close proximity to the first termination codon; thus, preventing detection of its abnormally long 3'UTR. It has been proposed that NMD

may have evolved to detect aberrant endogenous transcripts along with invading transcripts, such as those seen with an invading retrovirus or those produced by transposable elements. Thus, NMD may have evolved another strategy to identify aberrant transcripts that has not been identified.

The SV40 3'UTR is remarkably sensitive to NMD, such that mutations in genes that are absolutely required for NMD, including alleles of *Upf1* and *Upf2*, as well as genes that cause only a minor NMD defect, such as *Smg1* (Chen et al., 2005), are identifiable using a fluorescent reporter with the SV40 3'UTR (Metzstein and Krasnow, 2006). One hypothesis is that using only a part of the original viral genome in the SV40 3'UTR has separated it from any NMD insulating factors, similar to the RSV, allowing for efficient target recognition based on features like length and inefficient intron splicing.

Here, we analyze how the SV40 3'UTR is targeted by NMD in *Drosophila*. While we found the SV40 3'UTR is targeted by NMD, two features stimulate targeting including both length and intron; however, these may not be the only mechanisms involved in targeting the SV40 3'UTR. These data suggest that in contrast to the stability element in RSV, the SV40 3'UTR may contain a sensitizing element, like the downstream stability element (DSE) proposed in yeast (Peltz et al., 1993; Zhang et al., 1995), or the sensitizing sequence seen in immunoglobulin mu (Bühler et al., 2004).

We set out to determine what aspects of the SV40 3'UTR contribute to its NMD targeting by mapping features of the SV40 3'UTR that are required for NMD recognition. To this end, altered versions of the SV40 3'UTR were generated including constructs containing deletions or truncations, as well as a mutation in the splice donor of the inefficiently spliced intron, and constructs that completely lacked the intron. These

constructs contain an open reading frame encoding a GFP fluorescent reporter that are stably integrated into a unique site in the *Drosophila* genome; thus, providing a consistent level of transcription that allows expression levels of different constructs to be compared directly. Animals containing fluorescent reporter constructs harboring the altered SV40 3'UTRs were then examined in wild type and NMD loss of function backgrounds to determine features that were necessary for recognition of the SV40 3'UTR by NMD.

Materials and methods

Fly stocks

D. melanogaster lines were raised at 25°C on cornmeal/dextrose medium using standard conditions. The control chromosome used was *y w* FRT^{19A} (Xu and Rubin, 1993), the mutant chromosome used was *y w* FRT^{19A} $Upf2^{25G}$. The *e22c-GAL4* (Duffy et al., 1998; Takeshita et al., 2007) was used to drive expression of the reporters. The *breathless* promoter, *btl-Gal4*, used in combination with *UAS-GFP* as the reporter, (Shiga et al., 1996) was a secondary mechanism of expressing the reporter constructs. The *nlsDsRed2::SV40 3'UTR* reporter transgene was provided by M. Galko (M.D. Anderson Cancer Center). *e22c-GAL4, UAS-nlsDsRed2::SV40 3'UTR* was built using standard recombination methods (M.M.M., unpublished). Reporter constructs were injected into flies using ΦC31-mediated transgenesis (Venken et al., 2006) by Genetic Services Inc.

Cloning deletions and truncations of the SV40 3'UTR

To clone deletion and truncation constructs of the SV40 3'UTR, PCR primers were designed to amplify the SV40 3'UTR from a template *pUASTeGFP::SV40 3'UTR*

plasmid. Forward primers designed to amplify the 5' end of the SV40 3'UTR contained a fill sequence followed by an *XbaI* restriction site and a stop codon. Reverse primers near the 3' end of the SV40 3'UTR contained a fill sequence followed by a *StuI* restriction site. Primers were designed to make deletions and truncations within the SV40 3'UTR using a combination of outer restriction site-containing primers and inner primers with overlapping ends to zipper fragments of the SV40 3'UTR together. The SV40 3'UTR was amplified using phusion polymerase (NEB). Amplicons were imaged using an Alpha Innotech Gel Doc System and if a single band was present, the sample was treated with exo-SapIt (Affymetrix) or, if multiple bands were present, gel purified (Qiagen). Isolated amplicons were digested using *XbaI* and *StuI*, imaged using an Alpha Innotech Gel Doc System and gel purified (Qiagen). The *XbaI-StuI* digested products were ligated into the *XbaI-StuI* digested *pUAST attB* plasmid (a gift from Jyoti Misra, Thummel lab) and *UAS-eGFP* was cloned into the *KpnI* site. The *pUAST attB* plasmids were transformed into TOP10 cells (Invitrogen) and plasmids were isolated using a midi-prep (Qiagen). All constructs were sequenced and then sent to Genetic Services Inc. for injection into *attp16* (Venken et al., 2006) on the second chromosome.

Testing reporter enhancement

UAS-eGFP::SV40 3'UTR deletion and truncation males were mated to either *y w* FRT^{19A}*; e22c-GAL4, UAS-nlsDsRed2::SV40 3'UTR/CyO*, or *y w* FRT^{19A} $Upf2^{25G}$*; e22c-GAL4, UAS-nlsDsRed2::SV40 3'UTR/CyO* females. Wandering third instar larvae expressing both DsRed and GFP were observed using a Leica MZ16 fluorescence stereomicroscope, collected and frozen. Animals were also imaged using a Leica MZ16

fluorescence stereomicroscope and axiocam camera. A second round of collections was performed using a *Btl-GAL4* driver. For this experiment *y w FRT^{19A} Upf2^{25G}/FM7c or y w FRT^{19A}* females were mated to *Btl-GAL4 UAS-DsRed::SV40 3'UTR* males; the F1 progeny were mated to flies harboring *UAS-GFP::SV40 3'UTR* deletion and truncation reporters and were scored and collected as described above.

RNA isolation and quantitative RT-PCR

Total RNA was isolated from five to ten flash-frozen samples of adult males using Trizol reagent (Invitrogen) and phase-lock tubes (5-prime), and the RNeasy mini kit (Qiagen). Genomic contamination was reduced by on-column DNAse treatment (Ambion). RNA concentration was determined using a spectrophotometer and normalized for reverse transcription using the MuMulv reverse transcriptase Retroscript Kit (Ambion). Quantitative RT-PCR was performed on a BioRad iCycler thermocycler using SYBR green supermix (Bio-Rad). All experimental reactions were performed in triplicate with a single biological replicate. For quantitation samples were normalized to *RpL32*.

Results

To test what features of the SV40 3'UTR are required for targeting by NMD, twenty versions of the SV40 3'UTR were constructed. These constructs consisted of deletions and truncations of four sections of the SV40 3'UTR, which allowed us to test both length and sequence dependent detection of the SV40 3'UTR by NMD. For this experiment, samples were collected from two genotypes, a wild-type genotype where

NMD is functioning, along with an NMD mutant genotype, *Upf2*^{25G}, which is a hypomorphic allele that has a severe NMD defect (Avery et al., 2011; Frizzell et al., 2012; Metzstein and Krasnow, 2006). Animals were collected based on an internal control; a DsRed reporter with a full length SV40 3'UTR was expressed in the background of experimental *GFP::SV40 3'UTR* test reporters. Animals expressing the DsRed reporter at high intensity were collected in *Upf2*^{25G} mutants, while wild-type males expressing a reduced level of DsRed were collected. The different reporters were quantified using qRT-PCR with primers that detect either DsRed or GFP.

NMD targeting of the SV40 3'UTR is length dependent

To determine if NMD targeting of the SV40 3'UTR is dependent on the length of the SV40 3'UTR, a series of constructs were cloned that contain deletions and truncations of the SV40 3'UTR. Three sets of constructs were designed to also test whether the inefficiently spliced intron stimulates NMD independently of SV40 3'UTR length. The first set of constructs contain the sequence of the 66 bp inefficiently spliced intron; if the intron does somehow stimulate NMD, these constructs should be more sensitive to NMD than those not containing the intron. The second set of constructs contain the 66 bp sequence but have a mutation that changes the splice donor from GT to CC, which would presumably prevent splicing of the intron sequence, but would still provide the length of the 66 bp intron that would not be removed. Hence, these constructs could still theoretically stimulate NMD in a length dependent manner. The third set of constructs have the 66 bp intron sequence completely removed, so these would be predicted to be of

a shorter length and not undergo intron-based NMD stimulation, and thus might be less susceptible to NMD degradation.

All twenty constructs were tested for degree of NMD targeting by comparing the stability of the reporter in an NMD mutant to the stability in animals with functioning NMD. This analysis led to several observations that did not fit with our original predictions. First, the introduction of the GT-CC mutation in the splice donor site of the intron in the SV40 3'UTR completely destabilized the reporter (Fig. 5.1 A, B). This result corresponded with a low level or fluorescent protein expression seen in both wild type and NMD loss of function larvae. This mutation appears to have introduced sequence that destabilizes the reporter in an NMD-independent manner. No further analysis of these constructs was performed.

Next, it appears that overall, NMD targeting does depend on length, as the longer the SV40 3'UTR, the more efficiently it is targeted by NMD (Fig. 5.1 B). There does appear to be a length-dependent cutoff at 400 bp, below which no section of the SV40 3'UTR is capable of stimulating NMD-dependent destabilization. This is interesting because 400 bp is the average length of 3'UTRs in *Drosophila* and therefore may represent a general length-dependent threshold of NMD targeting in *Drosophila*. While these data support length-dependent targeting of the SV40 3'UTR, this interpretation cannot explain all the data, as there are several constructs that are not targeted by NMD even though others of the same length are targeted (Fig. 5.1 A, B). This phenomenon suggests there is some other feature of the SV40 3'UTR that is required for NMD-dependent degradation.

Intron in the SV40 3'UTR stimulates NMD

Analysis of the intron-containing and intron deleted constructs demonstrate that deleting the intron does not prevent the SV40 3'UTR from being targeted by NMD in the full length transcript (Fig. 5.1 A, B). This data agrees with a prior observation that deletion of the intron does not affect NMD targeting (Metzstein and Krasnow, 2006). However, the presence of the intron, or potentially the additional 66 bp left behind if the intron is not removed, does stimulate NMD targeting in the full-length constructs. However, the presence of the intron is not sufficient for targeting in the truncated constructs (Fig. 5.1 A, B). This finding is interesting as neither the presence of introns downstream of a termination codon, nor the presence of the exon junction complex components have been previously shown to stimulate NMD in *Drosophila* (Gatfield et al., 2003). Future studies will need to determine if the EJC is actually associated with the SV40 3'UTR, or if it is the presence of additional sequence that is causing this stimulation.

The 5' region of the SV40 3'UTR stimulates NMD targeting

To determine what other features of the SV40 3'UTR are required for NMD, I performed pairwise comparisons of the five segments of the SV40 3'UTR that were either deleted or truncated in the constructs (Fig. 5.2 A). Through these comparisons, I found the presence of the sequence in the 5' 3rd of the SV40 3'UTR (segments A and D in Figure 5.2 A, B) was the most likely region to stimulate NMD past what appears to be the length-dependent cut-off of ~400 bp. The presence of many of the segments in the 5' end of the SV40 3'UTR contribute to NMD targeting (Fig. 5.2 B, C, G, H). However, the

greatest stimulation came from having segments A and D (Fig. 5.2 B); these regions are the first segment after the termination codon and the section directly after the intron sequence. The presence of this 5' section provided even more stimulation of NMD targeting than the presence of the intron sequence (Fig. 5.2 C, F, I, K).

In contrast to the 5' end of the 3'UTR, the presence of the 3' 3rd of the SV40 3'UTR did not appear to have any effect on NMD targeting. The 3' (C segment in Figure 5.2 A, D, E, I, J) segment contains one of the two polyadenylation signals found in the SV40 3'UTR. Thus, I had predicted that disrupting this section, and therefore preventing polyadenylation, might alter the susceptibility of the SV40 3'UTR to NMD targeting. The observation that deleting this section had very little effect on NMD targeting suggests that it does not contain sequence important for NMD recognition, and suggests that the remaining polyadenylation signal is sufficient for proper polyadenylation of the reporter transcripts. The status of polyadenylation in these constructs could be tested in the future by performing 3' RACE on the different reporter transcripts to determine if they are indeed polyadenylated.

Discussion

This work describes features necessary for NMD targeting of an exogenous 3'UTR commonly used in *Drosophila* cloning constructs. Constructs were made that contained deletions and truncations of segments of the SV40 3'UTR to determine if there existed length, intron, or sequence specific features that render this 3'UTR an NMD target. We found that there is indeed a length-dependent threshold of about 400 bp, below which any region of the SV40 3'UTR is no longer an NMD target. The

confirmation that the SV40 3'UTR is targeted in a length-dependent manner was not particularly surprising as this is the proposed targeting mechanism utilized in *Drosophila*.

An observation that was surprising is that we did find some stimulation of NMD due to the presence of an inefficiently spliced intron in the SV40 3'UTR. The presence of an exon junction complex (EJC) deposited near the intron splice site is a feature that stimulates NMD recognition in plants (Kertész et al., 2006) and mammalian cells (Buchwald et al., 2010; Carter et al., 1996; Le Hir et al., 2001; Maquat, 2004; Zhang et al., 1998). However, this has not been shown before in *Drosophila* (Gatfield et al., 2003) and evidence even suggests that the EJC model of transcript targeting does not exist in this organism. One caveat to this observation is that using these particular constructs, I could not distinguish whether the EJC is associated with targeted SV40 3'UTR mRNA or if the unspliced intron sequence is extending the length of the 3'UTR, resulting in stimulation of length-dependent targeting. Unfortunately, the experiment that was originally designed to distinguish between length and EJC deposition did not work. A series of constructs were designed that contained a mutation, changing the GT-CC of the splice donor. These mutant transcripts would have retained the intron sequence and provided a direct test of length versus possible EJC deposition. However, the presence of this particular mutation completely destabilized the SV40 3'UTR in an NMD-independent manner, as there was no stabilization of these reporter transcripts in *Upf2*^{25G} mutant animals. Another option to test whether the added length of the intron stimulates NMD is to clone different filler sequence into the intron, which might avoid the destabilizing effects observed when the splice donor was mutated. Finally, a more direct test of the whether the EJC associates with the SV40 3'UTR involves cross-linking

immunoprecipitation of the EJC components and detection of associated mRNA sequences by qRT-PCR. However, this is a technically challenging experiment. While this experiment could determine if EJC-components interact with the SV30 3'UTR, it would not directly test whether EJC-bound mRNA is destabilized by NMD.

The most intriguing finding from mapping features of the SV40 3'UTR required for NMD targeting was a potential sequence-dependent targeting mechanism. Features near the 5' end of the SV40 3'UTR appear to be required for NMD targeting, whereas more 3' features appear to be dispensable for NMD targeting. The intron sequence is located in the 5' end of the SV40 3'UTR; however, the sequence directly after the stop codon and directly 3' of the intron sequence appear to be required for targeting, regardless of the presence or absence of the intron.

Experiments to follow up this observation should test whether the 5' sequence is sufficient for NMD targeting independent of length, which would suggest that there is sequence within the SV40 3'UTR that stimulates NMD. Since I was unable to make every combination of SV40 3'UTR rearrangement, it would be interesting in the future to see if just the 5' end of the SV40 3'UTR is sufficient for degradation by NMD. Combined, the A and D segments would result in a 258 bp 3'UTR, far below the threshold of the 400 bp length cut-off. Thus, any NMD degradation of this section would be extremely interesting and suggest there exists some sequence-dependent targeting that occurs in the SV40 3'UTR.

Interestingly, deleting the 3' 3rd of SV40 3'UTR appeared to have no effect on NMD targeting, suggesting that it does not contain sequence important for NMD recognition. This section contains one of the two polyadenylation signals, thus the

remaining polyadenylation signal may be sufficient for proper polyadenylation of the reporter transcripts. The status of polyadenylation in these constructs could be tested in the future by performing 3' RACE on the different reporter transcripts to determine if they are indeed polyadenylated.

A sequence-dependent targeting feature is interesting considering the Rous stability element discussed in the introduction. Sequence-dependent targeting could result from NMD evolving a mechanism to detect particular features of viral RNA that has itself evolved mechanisms to evade NMD. One hypothesis is that NMD may detect mRNA that looks more like a 3'UTR than an open reading frame. Sequence in 3'UTRs does not encode functional transcripts and tends to have more frequent stop codons in every reading frame. We looked at the number of stop codons found in each reading frame of the twenty constructs and found no difference in the proportion of stop codons that may make this sequence appear different to the NMD machinery. Another possibility is that the SV40 3'UTR does contain some sequence specific features that stimulate NMD. The downstream sequence element (DSE) was originally identified as a feature downstream of the termination codon that stimulates NMD (Peltz et al., 1993). While the concept of a consensus sequence stimulating NMD is intriguing and resembles other features known to direct targeting, like the AU-rich element (ARE) (Barreau et al., 2005), there is little evidence to suggest these are a common feature of NMD targets. Interestingly, the 5' end of the SV40 3'UTR contains four sequences that resemble the DSE consensus sequence identified in yeast (Table 5.1) (Zhang et al., 1995). In contrast, the 3' end of the SV40 3'UTR contains no sequence resembling the DSE. Further

analysis of the SV40 3'UTR will be interesting to see if there is some sequence-dependent NMD occurring in this 3'UTR.

In addition to mapping the *cis*-acting features required for targeting the SV40 3'UTR, the *trans*-acting NMD factors could also be tested for targeting efficiency. All of the NMD mutant analysis performed in this experiment was done in the $Upf2^{25G}$ mutant. These constructs could be tested in other NMD mutants to determine if multiple NMD pathways target this construct. Finally, all these experiments must be repeated as all graphs in this section are based on only one biological replicate of five to ten pooled larvae. The qRT-PCR data is also consistent with reporter fluorescence observed and imaged using two different GAL4 drivers, *e22c-GAL4* and *Btl-GAL4*, indicating the results described in this chapter will not vary drastically with the addition of more biological replicates.

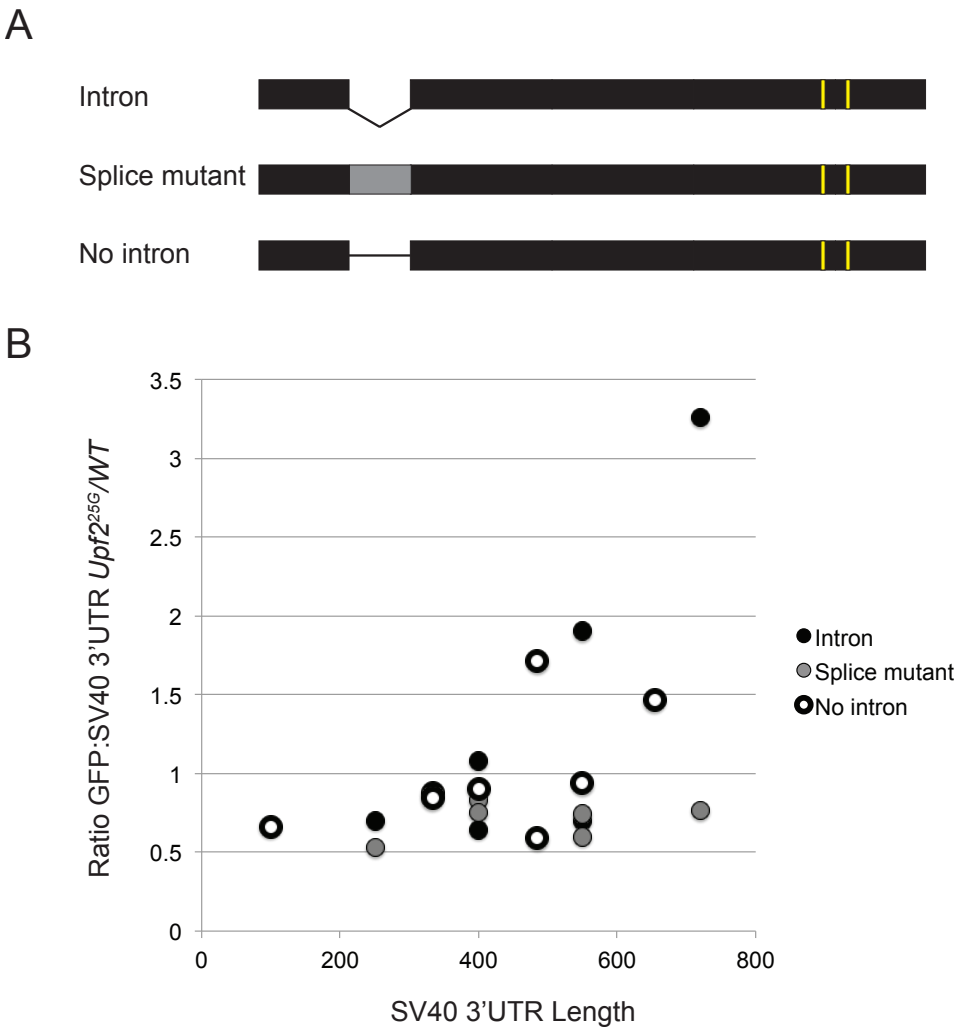


Figure 5.1 Length and intron-dependent stimulation of NMD targeting. (A) SV40 3'UTR constructs analyzed either contain an inefficiently spliced intron (Intron), contain a splice donor mutation GT-CC (Splice mutant), or the intron sequence has been deleted (no intron). Full-length constructs are shown. Yellow bars = polyadenylation sites. (B) Degree of NMD targeting determined by the ratio of GFP expression for each construct in *Upf2*^{25G} mutants over the expression in wild-type controls with functioning NMD. Values above one thus indicate the transcript is subject to NMD-dependent destabilization. The X-axis denotes the length of the transcript. Full length transcripts, are 720 bp, the original SV40 3'UTR (black circle, 720 bp) exhibits the maximum degree of NMD targeting. The full-length splice mutant (grey circle, 720 bp) was not subject to regulation by NMD. The full-length transcript with intron deleted (black open circle, 656 bp) exhibits a moderate degree of NMD targeting. Additional constructs containing the intron, splice mutant, or intron deletion are shown at decreasing lengths due to deletion of additional segments of the SV40 3'UTR. These exhibit a general length dependent targeting as the NMD-sensitivity decreases with SV40 3'UTR length. However, other features deleted within the SV40 3'UTR are required for NMD targeting as some transcripts fall below the targeting threshold of 1 on the Y-axis, but are longer than the 400 bp targeting threshold on the X-axis.

Figure 5.2 Features of the SV40 3'UTR required for NMD targeting. (A) SV40 3'UTR construct color-coded for sequences that mapped in the different SV40 3'UTR constructs. Five features were analyzed, including the presence or absence of intron sequence E, green segment. Yellow bars = polyadenylation sites. (B-K) Degree of NMD targeting determined by the ratio of GFP expression for each construct in *Upf2*^{25G} mutants over the expression in wild-type controls with functioning NMD. Data points with values above one are subject to NMD-dependent destabilization. The X-axis denotes length of the transcript, which varies based on the segment (SV40 3'UTR A-E) deleted in the construct. For example, the 250 bp contains segments D and E. Each graph represents the NMD-targeting of all the examined constructs, however, the colors indicate the inclusion of that segment in the construct. (B) The combination of A and D segments stimulates NMD. Constructs containing both the A and D segments are in blue-purple stripes. Constructs containing only the A segment are purple, only the D segment are blue, constructs that contain neither A or D are grey. The differences in length are due to deletion of different segments of the SV40 3'UTR. The combination of A and D stimulates NMD targeting as all constructs containing A and D above 400 bp in length are subject to NMD (above one on the Y-axis). This stimulation is most apparent at 550 bp, where the constructs containing only A (purple) or only D (blue) are below the NMD target threshold of one on the Y-axis but are above the length threshold. However, the combination of A and D (blue-purple stripes) stimulates NMD targeting. (C) The combination of A and E (intron) can stimulate NMD targeting, but the combination of both segments is not required for targeting and E is never sufficient on its own for targeting. Constructs containing only A are purple, those containing only E are green and those containing both are green and purple striped. Constructs containing both A and E can stimulate targeting above the 400 bp length threshold, but this is dependent on the presence of segment A. No segments that contain only E (the intron) without the A segment are targeted by NMD. Perhaps the inclusion of the A segment is required for efficient splicing of E. (D) The C segment does not contribute to NMD targeting. The A segment alone (purple) or the C segment alone (red) or the combination of the two (red-purple stripes) shows that the C segment on its own, even at 550 bp (above the 400 bp length threshold) is not sufficient to cause NMD targeting. The only segments the contain C that are targeted by NMD also contain the A segment. However, the A segment can be targeted without the presence of the C segment (see purple data points at 480 bp and 550 bp). (E) Neither B or C, nor the combination of B and C stimulate NMD targeting. The combination of B and C (red-orange stripes) falls below the targeting threshold of 1 on the Y-axis in most transcripts above 400 bp, so these sections likely do not contain features targeted by NMD. (F) The combination of D (blue) and E (green) does not show a pattern of NMD targeting. No constructs of only E were made. These two segments are the smallest in the series, so even length dependent stimulation of these segments will be difficult to determine. (G) The combination of the A and B segments may stimulate NMD targeting. The combination of the A and B segments (orange-purple striped) tends to be targeted by NMD; however, at 400 bp the A segment alone (purple) is targeted, whereas the B segment alone (orange) is not, suggesting A may contribute more important sequence to NMD targeting than B. (H) The combination of B (orange) and D (red) does not result in a pattern indicative of one or the combination of these two segments stimulating NMD targeting in these constructs. (I) The C and E segments do not contribute to NMD

targeting together. The combination of C (red) with E (green) actually seems to reduce NMD targeting of these constructs between 400-550 bp (red-green stripes). (J) The C and D segments do not contribute to NMD targeting together. The combination of C (red) with D (blue) actually seems to reduce NMD targeting of these constructs between 400-550 bp (red-blue stripes). (K) The combination of B (orange) and E (green) may stimulate NMD in longer constructs, but E is not sufficient for targeting in the longer segments. Grey data points = features not present in the constructs examined.

Figure 5.2 Continued.

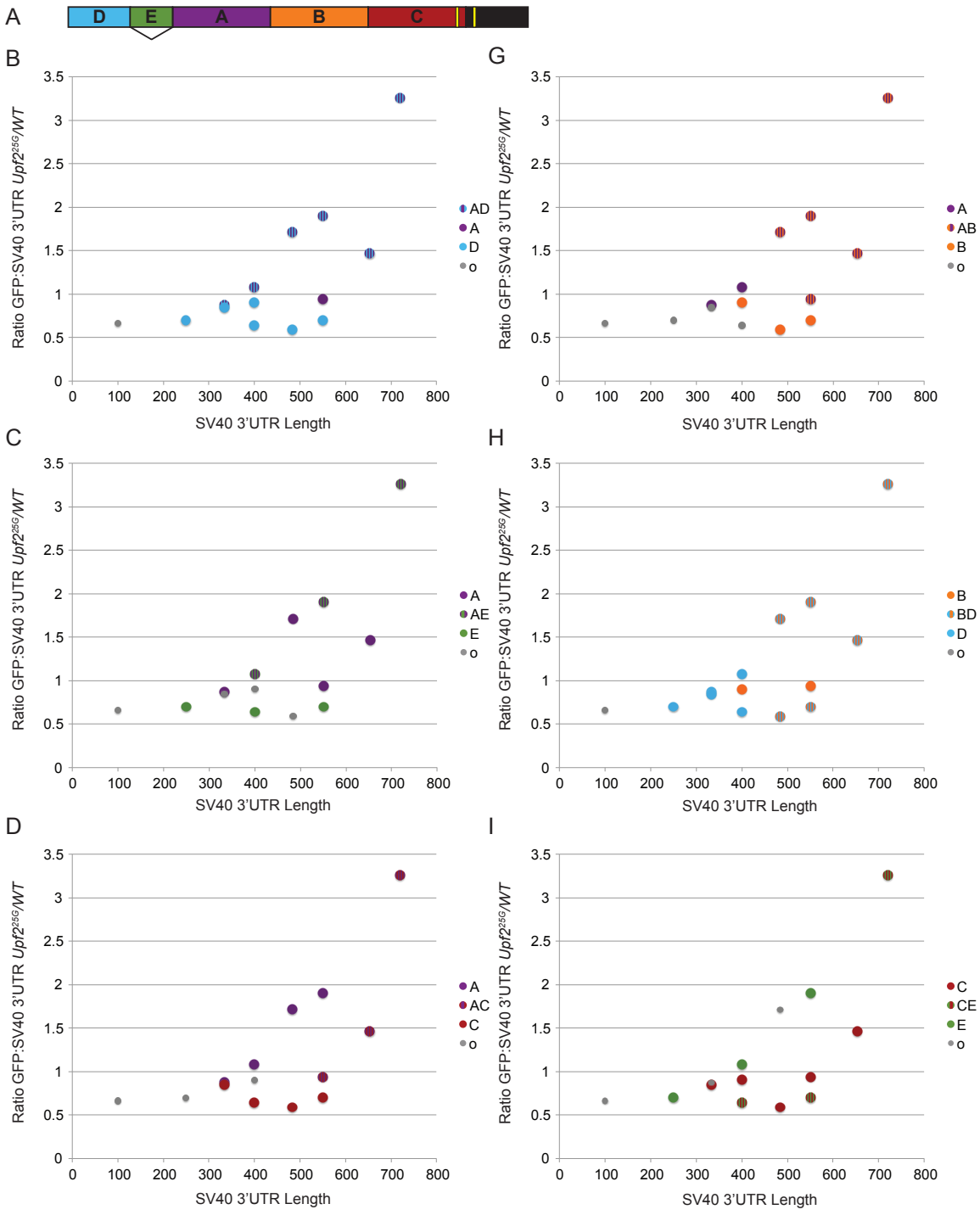


Figure 5.2 Continued

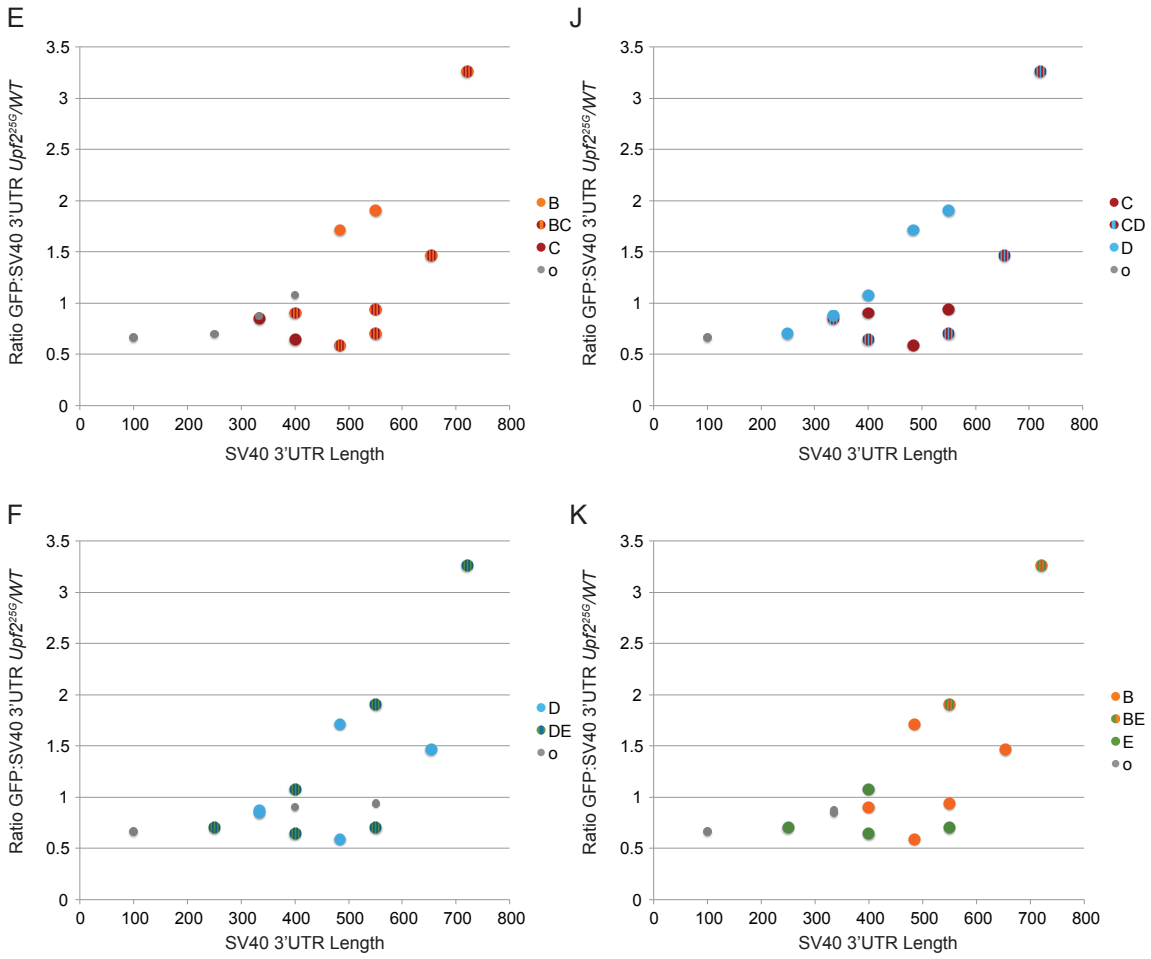


Table 5.1 DSE Consensus sequences in the SV40 3'UTR

DSE Motif[1]		**T**	**G**	**Y**	**Y**	**G**	**A**	**T**	**G**	**Y**	**Y**	**Y**	**Y**
SV40 3'UTR Segment[2]	A	A	A	**C**	**T**	**G**	**A**	**T**	**G**	A	A	**T**	G
	A	**T**	A	G	**T**	**G**	**A**	**T**	**G**	A	**T**	G	A
	A	**T**	**G**	A	**T**	**G**	**A**	**T**	**G**	A	G	G	**C**
	B	A	T	**T**	**T**	**G**	**A**	**T**	**G**	**T**	A	**T**	A

[1] Consensus sequence found in (Zhang et al., 1995). Y = C or T, Bold = match between SV40 3'UTR sequence and consensus.
[2] A segment (Fig. 5.2) sequence 3' to the intron, B segment (Fig. 5.2) sequence in the middle of the SV40 3' UTR.

References

Amrani, N., Ganesan, R., Kervestin, S., Mangus, D.A., Ghosh, S., and Jacobson, A. (2004). A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature *432*, 112–118.

Avery, P., Vicente-Crespo, M., Francis, D., Nashchekina, O., Alonso, C.R., and Palacios, I.M. (2011). Drosophila Upf1 and Upf2 loss of function inhibits cell growth and causes animal death in a Upf3-independent manner. Rna *17*, 624–638.

Barreau, C., Paillard, L., and Osborne, H.B. (2005). AU-rich elements and associated factors: are there unifying principles? Nucl. Acids Res. *33*, 7138–7150.

Buchwald, G., Ebert, J., Basquin, C., Sauliere, J., Jayachandran, U., Bono, F., Le Hir, H., and Conti, E. (2010). Insights into the recruitment of the NMD machinery from the crystal structure of a core EJC-UPF3b complex. Proc. Natl. Acad. Sci. USA *107*, 10050–10055.

Bühler, M., Paillusson, A., and Mühlemann, O. (2004). Efficient downregulation of immunoglobulin mu mRNA with premature translation-termination codons requires the 5'-half of the VDJ exon. Nucl. Acids Res. *32*, 3304–3315.

Carter, M.S., Li, S., and Wilkinson, M.F. (1996). A splicing-dependent regulatory mechanism that detects translation signals. Embo J. *15*, 5965–5975.

Celniker, S.E., and Rubin, G.M. (2003). The Drosophila melanogaster genome. Annu. Rev. Genomics Hum. Genet. *4*, 89–117.

Chen, Z., Smith, K.R., Batterham, P., and Robin, C. (2005). Smg1 nonsense mutations do not abolish nonsense-mediated mRNA decay in Drosophila melanogaster. Genetics *171*, 403–406.

Duffy, J.B., Harrison, D.A., and Perrimon, N. (1998). Identifying loci required for follicular patterning using directed mosaics. Development *125*, 2263–2271.

Eberle, A.B., Stalder, L., Mathys, H., Zamudio Orozco, R., and Mühlemann, O. (2008). Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. PLoS Biol. *6*, e92.

Frizzell, K.A., Rynearson, S.G., and Metzstein, M.M. (2012). Drosophila mutants show NMD pathway activity is reduced, but not eliminated, in the absence of Smg6. Rna *18*, 1475–1486.

Gatfield, D., Unterholzner, L., Ciccarelli, F.D., Bork, P., and Izaurralde, E. (2003). Nonsense-mediated mRNA decay in Drosophila: at the intersection of the yeast and mammalian pathways. Embo J. *22*, 3960–3970.

Hogg, J.R., and Goff, S.P. (2010). Upf1 Senses 3'UTR Length to Potentiate mRNA

Decay. Cell *143*, 379–389.

Kerényi, Z., Mérai, Z., Hiripi, L., Benkovics, A.H., Gyula, P., Lacomme, C., Barta, E., Nagy, F., and Silhavy, D. (2008). Inter-kingdom conservation of mechanism of nonsense-mediated mRNA decay. Embo J. *27*, 1585–1595.

Kertész, S., Kerényi, Z., Mérai, Z., Bartos, I., Pálfy, T., Barta, E., and Silhavy, D. (2006). Both introns and long 3'-UTRs operate as cis-acting elements to trigger nonsense-mediated decay in plants. Nucl. Acids Res. *34*, 6147–6157.

Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M.J. (2001). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. Embo J. *20*, 4987–4997.

Maquat, L.E. (2004). Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. Nat. Rev. Mol. Cell Biol. *5*, 89–99.

Metzstein, M.M., and Krasnow, M.A. (2006). Functions of the nonsense-mediated mRNA decay pathway in Drosophila development. PLoS Genet. *2*, e180.

Muhlrad, D., and Parker, R. (1999). Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance. Rna *5*, 1299–1307.

Peltz, S.W., Brown, A.H., and Jacobson, A. (1993). mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor. Genes Dev. *7*, 1737–1754.

Rehwinkel, J., Letunic, I., Raes, J., Bork, P., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. Rna *11*, 1530–1544.

Shiga, Y., Tanaka-Matakatsu, M., and Hayashi, S. (1996). A nuclear GFP/ß-galactosidase fusion protein as a marker for morphogenesis in living Drosophila. Develop. Growth Differ. 99–106.

Singh, G., Rebbapragada, I., and Lykke-Andersen, J. (2008). A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. PLoS Biol. *6*, e111.

Takeshita, D., Zenno, S., Lee, W.C., Saigo, K., and Tanokura, M. (2007). Crystal structure of the PIN domain of human telomerase-associated protein EST1A. Proteins *68*, 980–989.

Thummel, C.S., Boulet, A.M., and Lipshitz, H.D. (1988). Vectors for Drosophila P-element-mediated transformation and tissue culture transfection. Gene *74*, 445–456.

Venken, K.J.T., He, Y., Hoskins, R.A., and Bellen, H.J. (2006). P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster.

Science *314*, 1747–1751.

Withers, J.B., and Beemon, K.L. (2010a). Structural features in the Rous sarcoma virus RNA stability element are necessary for sensing the correct termination codon. Retrovirology *7*, 65.

Withers, J.B., and Beemon, K.L. (2010b). Structural features in the Rous sarcoma virus RNA stability element are necessary for sensing the correct termination codon. Retrovirology *7*, 65.

Withers, J.B., and Beemon, K.L. (2011). The structure and function of the rous sarcoma virus RNA stability element. J. Cell. Biochem. *112*, 3085–3092.

Wittkopp, N., Huntzinger, E., Weiler, C., Sauliere, J., Schmidt, S., Sonawane, M., and Izaurralde, E. (2009). Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic development and survival. Mol. Cell. Biol. *29*, 3517–3528.

Xu, T., and Rubin, G.M. (1993). Analysis of genetic mosaics in developing and adult Drosophila tissues. Development *117*, 1223–1237.

Zhang, J., Sun, X., Qian, Y., LaDuca, J.P., and Maquat, L.E. (1998). At least one intron is required for the nonsense-mediated decay of triosephosphate isomerase mRNA: a possible link between nuclear splicing and cytoplasmic translation. Mol. Cell. Biol. *18*, 5272–5283.

Zhang, S., Ruiz-Echevarría, M.J., Quan, Y., and Peltz, S.W. (1995). Identification and characterization of a sequence motif involved in nonsense-mediated mRNA decay. Mol. Cell. Biol. *15*, 2231–2244.

CHAPTER 6

CONCLUSIONS

Nonsense mediated decay (NMD) is a cellular surveillance mechanism originally identified for its role in clearing mRNAs encoding premature termination codons (PTCs) from the cell (Chang and Kan, 1979; Hodgkin et al., 1989; Leeds et al., 1991; Losson and Lacroute, 1979). More recent efforts have shown that a large proportion, up to 10% of the genome in many organisms, of endogenous transcripts are also regulated by NMD (He et al., 2003; Mendell et al., 2004; Rehwinkel et al., 2005; Weischenfeldt et al., 2008; Wittmann et al., 2006). This suggests NMD plays a role in development and physiology. Thus, it is important to perform *in vivo* analysis of nonsense mediated decay.

The work presented in this thesis focused on three broad questions relevant to the mechanism of NMD. First, have all the genes required for NMD been identified? Second, what are the roles of NMD pathway genes *in vivo*? Third, how does NMD differentiate a native termination codon from a premature termination codon?

Chapter 2 and Chapter 3 describe work to address whether all genes required for NMD have been identified. Chapter 2 presents work performed for the first forward genetic screen for genes involved in NMD in *Drosophila* (Frizzell et al., 2012). Using a mosaic approach we identified both viable and lethal alleles of previously known NMD genes; including alleles of *Upf1*, *Upf2*, *Smg1,* and *Smg6*. While alleles of *Upf1*, *Upf2*,

and *Smg1* had previously been analyzed in *Drosophila* (Chen et al., 2005; Metzstein and Krasnow, 2006), Chapter 2 provides the first mutant analysis of the NMD gene, *Smg6* (Frizzell et al., 2012). Through this work we were able to address a key question in the NMD field: what roles do the NMD pathway genes play *in vivo*? Our mutant analysis of *Smg6* resulted in two surprising observations. First, alleles of *Smg6* are viable, (albeit, not fully) in complete loss-of-function mutants. The viability of *Smg6* was surprising considering loss of *Smg6* function may result in lethality in *C. elegans* (Cali and Anderson, 1998). Additionally, zebrafish morpholino data indicates *Smg6* knockdown results in more severe developmental defects than knockdown of the core NMD component, *Upf1* (Wittkopp et al., 2009). Our results could show that *Smg6* is less essential for viability in *Drosophila*, suggesting that additional lethal mutations near the *Smg6* locus and morpholino off-target effects could have caused the lethality seen in *C. elegans* and zebrafish, respectively.

The second surprising result is loss of *Smg6* in *Drosophila* does not result in a severe NMD defect. Smg6 is an endonuclease that cleaves targeted mRNAs near the PTC (Eberle et al., 2009; Gatfield and Izaurralde, 2004). In *Drosophila* S2 cells endonucleolytic cleavage was the only mechanism observed to free the ends of the mRNA for the rapid degradation that is characteristic of NMD (Huntzinger et al., 2008). Our *in vivo* results suggest other degradation mechanisms are utilized in *Drosophila*. Indeed, in addition to endonucleolytic cleavage by *Smg6*, NMD is also known to result in deadenylation-independent decapping followed by degradation by the exonuclease, XRN1 (*pacman* in *Drosophila*), and to a lesser extent, rapid deadenylation followed by degradation by the exosome in mammalian cells (Chen and Shyu, 2003; Couttet and

Grange, 2004; Lejeune et al., 2003) and yeast (He and Jacobson, 2001). Perhaps these alternative mechanisms are used *in vivo* in *Drosophila*, but are absent in S2 cells, hence the lack of detection of these mechanisms in previous studies (Gatfield and Izaurralde, 2004; Huntzinger et al., 2008).

One final observation resulting from the analysis of *Smg6* involves the role of Smg1 in NMD. Smg1 is a phosphatidylinositol kinase-like kinase (PIKK) that phosphorylates Upf1 (Denning et al., 2001; Yamashita et al., 2001), providing binding sites for NMD components, including Smg6 (Okada-Katsuhata et al., 2011). Loss of *Smg1* in *Drosophila* results in minor NMD defects (Chen et al., 2005; Metzstein and Krasnow, 2006). However, it is thought that phosphorylation of Upf1 by Smg1 is essential for Smg6 recruitment to the targeted transcript. Surprisingly, *Smg6* mutants have a more severe NMD defect than *Smg1* mutants (Frizzell et al., 2012). This result suggests that either Smg6 is recruited to the complex independent of the Smg1 phosphorylation of Upf1, or alternatively, another kinase may be capable of phosphorylating Upf1, thus providing the binding platform required for NMD recruitment. Our data from analysis of *Smg6* provides preliminary evidence that more genes, potentially other nucleases and kinases, may be involved in NMD *in vivo*. Thus, alleles of *Upf1*, *Upf2*, *Smg1*, and *Smg6* identified in our screens have provided valuable insight into the possibility of additional genes or pathways utilized for NMD dependent degradation, and will provide the community with valuable reagents for further study of NMD in *Drosophila*.

In addition to alleles of known NMD genes, our screen resulted in the isolation of seven new lines that could be new NMD genes. Chapter 3 describes preliminary

mapping data and analysis of these seven lines. Five lines make up a single lethal complementation group that was mapped to CG33970, a previously uncharacterized gene. CG33970 encodes a putative ABC-transporter that is conserved in invertebrates (Dean and Annilo, 2005) and has sequence homology to a ribosome-bound ABC-transporter, RbbA, that may function in translation termination in *E. coli* (Kiel and Ganoza, 2001; Kiel et al., 1999; Xu et al., 2006). However, evidence described in Chapter 3 suggests CG33970 may not be the gene responsible for the NMD defect, but instead is only responsible for the lethality associated with these five lines. Further work should be done to verify the role of CG33970 to determine if it really is a novel NMD gene. Two additional lines isolated in the screen on 3R complement alleles in the lethal complementation group, and complement each other. Mapping and characterization of these two lines may result in the identification of novel NMD genes.

Chapter 4 of this thesis provides a method that utilizes alleles of *Smg6* and other NMD components to address the role of NMD *in vivo*. In this chapter, I begin to characterize the endonucleolytic cleavage products produced after NMD targeting, and (presumably), Smg6 cleavage. This analysis utilizes RNAi knockdown of the 5' exonuclease *pacman* (*pcm*), allowing for detection of the 3' cleavage products whose 5' end would be free to degradation by *pcm* after endonucleolytic cleavage by Smg6 (Chang and Kan, 1979; Franks et al., 2010; Hodgkin et al., 1989; Leeds et al., 1991; Losson and Lacroute, 1979; Rehwinkel et al., 2005). Smg6 is known to cleave NMD-targeted mRNAs near the PTC (Eberle et al., 2009; Gatfield and Izaurralde, 2004; He et al., 2003; Huntzinger et al., 2008; Mendell et al., 2004; Rehwinkel et al., 2005; Weischenfeldt et al., 2008; Wittmann et al., 2006), however this has only been shown in tissue culture. *In vivo*,

this assay could provide a way to characterize the different defects associated with *Smg6* hypomorphic alleles. For instance, a hypomorphic allele, $Smg6^{2a}$, contains a missense mutation predicted to disrupt endonuclease function (Frizzell et al., 2012; Glavan et al., 2006; Takeshita et al., 2007). However, we find this allele appears to have some residual endonuclease activity, suggesting that the missense mutation does not fully disrupt the endonuclease domain as predicted (Chen et al., 2005; Glavan et al., 2006; Metzstein and Krasnow, 2006; Takeshita et al., 2007). Additionally, this type of cleavage assay can be used to determine whether a transcript is directly or indirectly targeted by NMD (Frizzell et al., 2012; Rehwinkel et al., 2005). An indirect target may be stabilized upon loss of NMD function, but will not show endonuclease-specific stabilization of the 3' end of the transcript in a *pcm* knockdown. However, a direct target should result in 3' fragment stabilization using this assay. Transcript fragment specific stabilization is a read-out of the endonuclease-specific cleavage by Smg6 and both 5' and 3' fragments should be stabilized in a *Smg6* mutant. While work must be done to optimize this assay, there is potential for interesting results regarding the cleavage characteristics of Smg6. Additionally, characterizing other NMD genes using this assay may provide insight as to the specific mechanism of NMD action. A key observation was that a strong hypomorphic allele of *Upf2*, $Upf2^{25G}$, resulted in 3' fragment stabilization of one NMD target and 5' stabilization of another NMD target. One hypothesis is that stabilization resulted from transcript protection by an NMD mRNP that was locked on the mRNA due to the nature of the mutation in $Upf2^{25G}$ (Cali and Anderson, 1998; Metzstein and Krasnow, 2006). While this is currently speculative, it could represent an interesting

avenue of investigation into the composition and location of the NMD mRNP on targeted transcripts (Franks et al., 2010; Wittkopp et al., 2009).

Finally, Chapter 5 of this thesis addresses the question of how a premature termination codon is differentiated from a natural termination codon. This chapter describes mapping of features of the exogenous SV40 3'UTR, that are required for targeting by NMD (Eberle et al., 2009; Gatfield and Izaurralde, 2004; Metzstein and Krasnow, 2006). Three features were determined to stimulate NMD targeting of the SV40 3'UTR. The first feature was length; the full length (720 bp) SV40 3'UTR is targeted by NMD. Deleting sections of the SV40 3'UTR resulted in NMD targeting until the length dropped below a threshold of ~400 bp. This result correlates with the current length-dependent model of NMD targeting. The length model posits that longer 3'UTRs or PTCs are subject to NMD because translation termination at an early termination codon appears as a long 3'UTR due to inefficient interactions between eRF3 and poly-A-binding proteins associated with the poly-A-tail. Thus, allowing Upf1 to compete with poly-A-binding protein for an interaction with eRF3. Upf1 association with eRF3 then initiates assembly of the SURF complex, which triggers NMD (Amrani et al., 2004; Behm-Ansmant et al., 2007; Eberle et al., 2008; Hogg and Goff, 2010; Huntzinger et al., 2008; Kashima et al., 2006; Muhlrad and Parker, 1999; Singh et al., 2008). The length-dependent mechanism is likely utilized for targeting the SV40 3'UTR for degradation as NMD targeting is relieved upon truncation of the SV40 3'UTR. Length, however, cannot be the only method utilized for transcript targeting, as several of the reporters are larger than 400 bp and are not NMD substrates. While the SV40 3'UTR is considered long for an NMD target in *Drosophila*, where the average 3'UTR length is ~400 bp, there must be

other targeting features as there is little correlation between 3'UTR length and targeting of endogenous transcripts in NMD mutants (A. Chapin, personal communication) (Barberan-Soler et al., 2009; Chen and Shyu, 2003; Couttet and Grange, 2004; Frizzell et al., 2012; Lejeune et al., 2003; Lewis et al., 2003; Metzstein and Krasnow, 2006; Nguyen et al., 2012; Ramani et al., 2009; Rehwinkel et al., 2005).

The second feature that stimulates NMD targeting of the SV40 3'UTR is the presence of an intron. A proposed model of NMD targeting depends on the presence of an intron downstream of a premature termination codon. Upon splicing and removal of the intron sequence, a protein complex called the exon junction complex (EJC) is deposited on the mRNA 20-24 nt upstream of the splice site (He and Jacobson, 2001; Le Hir et al., 2000). The EJC provides a binding site for Upf3, a component of the NMD machinery (Gatfield and Izaurralde, 2004; Huntzinger et al., 2008; Le Hir et al., 2001). The EJC is normally removed or remodeled during translation. However, if the ribosome terminates prior to the last splice site or if there is an intron found in the 3'UTR, the EJC remains associated with the message, which triggers NMD (Buchwald et al., 2010; Denning et al., 2001; Ivanov et al., 2008; Singh et al., 2008; Yamashita et al., 2001). While the EJC has been shown to stimulate NMD in plants (Kerényi et al., 2008; Kertész et al., 2006; Okada-Katsuhata et al., 2011) and mammalian cells (Chen et al., 2005; Isken and Maquat, 2007; Metzstein and Krasnow, 2006), there is currently no evidence showing NMD is stimulated by an EJC in *Drosophila* (Frizzell et al., 2012; Gatfield et al., 2003). Data in Chapter 5 suggest the presence of an intron in the SV40 3'UTR stimulates NMD. If this is due to the presence of an EJC associated with the SV40 3'UTR, it would be the first time this mechanism of target recognition has been shown in *Drosophila*.

One caveat to this result is that the intron could remain intact, thus providing extra sequence and length to the 3'UTR by introducing confusion with the length-dependent model of NMD. The status of the intron in the SV40 3'UTR needs to be determined before a final conclusion can be made on this observation.

The final feature determined to stimulate NMD targeting the SV40 3'UTR is the presence of a specific sequence in the 5' end of the SV40 3'UTR. This observation is reminiscent of a feature identified in yeast, termed the downstream sequence element (DSE) (Dean and Annilo, 2005; Peltz et al., 1993; Zhang et al., 1995). Indeed, much of the consensus sequence identified in the DSE (Kiel and Ganoza, 2001; Kiel et al., 1999; Xu et al., 2006; Zhang et al., 1995) is found in the SV40 3'UTR, but only in the segments constituting the 5' end of the 3' UTR. A similar, sequence dependent result was also observed in mammalian cells, where a 177 bp 5' region of the immunoglobulin mu transcript was shown to be required for efficient NMD (Bühler et al., 2004). This data suggests that some sequence-dependent NMD may occur, whether this is due to direct binding of RNA-binding proteins to a particular sequence or to secondary structure that results from the particular sequence is not known. However, this result does suggest the SV40 3'UTR could be used to determine the minimal SV40 3'UTR sequence sufficient for NMD targeting, and this could be very interesting if that sequence contained that putative DSE.

In conclusion, work presented in this thesis suggests several possibilities for NMD in *Drosophila*. First, analysis of alleles recovered in the X and 3R screens suggest more genes are likely required for NMD in *Drosophila*. These screens only covered ~40% of the genome and the 3R screen is not yet saturated. Second, further work on

these alleles will be important for addressing whether lethality associated with loss of particular NMD genes is due to loss of NMD function or non-NMD roles of those genes. Data presented here suggest that the severity of NMD defects observed with loss of *Smg6*, *Upf2*, or *Smg1* correlate with the severity of the developmental defects associated with loss of those genes. It will be interesting to see if this pattern persists with other known NMD genes. Third, additional tests that can differentiate between direct and indirect NMD targets, like the *pcm* knockdown assay described in Chapter 4, will be useful for further characterization of these alleles and the transcripts targeted by NMD *in vivo.* Future work characterizing cleavage and mRNP composition are ways to begin assaying mechanistic features of NMD *in vivo*. Finally, work in Chapter 5 suggests that in addition to a length-dependent targeting mechanism, *Drosophila* may also utilize intron-dependent and sequence-dependent targeting in particular cases. Future experiments are needed to validate these observations, but these data could present a new way of thinking about NMD *in vivo*.

References

Amrani, N., Ganesan, R., Kervestin, S., Mangus, D.A., Ghosh, S., and Jacobson, A. (2004). A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature *432*, 112–118.

Barberan-Soler, S., Lambert, N.J., and Zahler, A.M. (2009). Global analysis of alternative splicing uncovers developmental regulation of nonsense-mediated decay in C. elegans. Rna *15*, 1652–1660.

Behm-Ansmant, I., Gatfield, D., Rehwinkel, J., Hilgers, V., and Izaurralde, E. (2007). A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. Embo J. *26*, 1591–1601.

Buchwald, G., Ebert, J., Basquin, C., Sauliere, J., Jayachandran, U., Bono, F., Le Hir, H., and Conti, E. (2010). Insights into the recruitment of the NMD machinery from the crystal structure of a core EJC-UPF3b complex. Proc. Natl. Acad. Sci. USA *107*, 10050–10055.

Bühler, M., Paillusson, A., and Mühlemann, O. (2004). Efficient downregulation of immunoglobulin mu mRNA with premature translation-termination codons requires the 5'-half of the VDJ exon. Nucl. Acids Res. *32*, 3304–3315.

Cali, B.M., and Anderson, P. (1998). mRNA surveillance mitigates genetic dominance in Caenorhabditis elegans. Mol. Gen. Genet. *260*, 176–184.

Chang, J.C., and Kan, Y.W. (1979). beta 0 thalassemia, a nonsense mutation in man. Proc. Natl. Acad. Sci. USA *76*, 2886–2889.

Chen, C.-Y.A., and Shyu, A.-B. (2003). Rapid deadenylation triggered by a nonsense codon precedes decay of the RNA body in a mammalian cytoplasmic nonsense-mediated decay pathway. Mol. Cell. Biol. *23*, 4805–4813.

Chen, Z., Smith, K.R., Batterham, P., and Robin, C. (2005). Smg1 nonsense mutations do not abolish nonsense-mediated mRNA decay in Drosophila melanogaster. Genetics *171*, 403–406.

Couttet, P., and Grange, T. (2004). Premature termination codons enhance mRNA decapping in human cells. Nucl. Acids Res. *32*, 488–494.

Dean, M., and Annilo, T. (2005). Evolution of the ATP-binding cassette (ABC) transporter superfamily in vertebrates. Annu. Rev. Genomics Hum. Genet. *6*, 123–142.

Denning, G., Jamieson, L., Maquat, L.E., Thompson, E.A., and Fields, A.P. (2001). Cloning of a novel phosphatidylinositol kinase-related kinase: characterization of the human SMG-1 RNA surveillance protein. J. Biol. Chem. *276*, 22709–22714.

Eberle, A.B., Lykke-Andersen, S., Mühlemann, O., and Jensen, T.H. (2009). SMG6

promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat. Struct. Mol. Biol. *16*, 49–55.

Eberle, A.B., Stalder, L., Mathys, H., Zamudio Orozco, R., and Mühlemann, O. (2008). Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. PLoS Biol. *6*, e92.

Franks, T.M., Singh, G., and Lykke-Andersen, J. (2010). Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense-mediated mRNA decay. Cell *143*, 938–950.

Frizzell, K.A., Rynearson, S.G., and Metzstein, M.M. (2012). Drosophila mutants show NMD pathway activity is reduced, but not eliminated, in the absence of Smg6. Rna *18*, 1475–1486.

Gatfield, D., and Izaurralde, E. (2004). Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. Nature *429*, 575–578.

Gatfield, D., Unterholzner, L., Ciccarelli, F.D., Bork, P., and Izaurralde, E. (2003). Nonsense-mediated mRNA decay in Drosophila: at the intersection of the yeast and mammalian pathways. Embo J. *22*, 3960–3970.

Glavan, F., Behm-Ansmant, I., Izaurralde, E., and Conti, E. (2006). Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. Embo J. *25*, 5117–5125.

He, F., and Jacobson, A. (2001). Upf1p, Nmd2p, and Upf3p regulate the decapping and exonucleolytic degradation of both nonsense-containing mRNAs and wild-type mRNAs. Mol. Cell. Biol. *21*, 1515–1530.

He, F., Li, X., Spatrick, P., Casillo, R., Dong, S., and Jacobson, A. (2003). Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5“ to 3” mRNA decay pathways in yeast. Mol. Cell *12*, 1439–1452.

Hodgkin, J., Papp, A., Pulak, R., Ambros, V., and Anderson, P. (1989). A new kind of informational suppression in the nematode Caenorhabditis elegans. Genetics *123*, 301–313.

Hogg, J.R., and Goff, S.P. (2010). Upf1 Senses 3'UTR Length to Potentiate mRNA Decay. Cell *143*, 379–389.

Huntzinger, E., Kashima, I., Fauser, M., Saulière, J., and Izaurralde, E. (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. Rna *14*, 2609–2617.

Isken, O., and Maquat, L.E. (2007). Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. Genes Dev. *21*, 1833–1856.

Ivanov, P.V., Gehring, N.H., Kunz, J.B., Hentze, M.W., and Kulozik, A.E. (2008). Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. Embo J. *27*, 736–747.

Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G., and Ohno, S. (2006). Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. Genes Dev. *20*, 355–367.

Kerényi, Z., Mérai, Z., Hiripi, L., Benkovics, A.H., Gyula, P., Lacomme, C., Barta, E., Nagy, F., and Silhavy, D. (2008). Inter-kingdom conservation of mechanism of nonsense-mediated mRNA decay. Embo J. *27*, 1585–1595.

Kertész, S., Kerényi, Z., Mérai, Z., Bartos, I., Pálfy, T., Barta, E., and Silhavy, D. (2006). Both introns and long 3'-UTRs operate as cis-acting elements to trigger nonsense-mediated decay in plants. Nucl. Acids Res. *34*, 6147–6157.

Kiel, M.C., and Ganoza, M.C. (2001). Functional interactions of an Escherichia coli ribosomal ATPase. Eur. J. Biochem. *268*, 278–286.

Kiel, M.C., Aoki, H., and Ganoza, M.C. (1999). Identification of a ribosomal ATPase in Escherichia coli cells. Biochimie *81*, 1097–1108.

Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M.J. (2001). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. Embo J. *20*, 4987–4997.

Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. (2000). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. Embo J. *19*, 6860–6869.

Leeds, P., Peltz, S.W., Jacobson, A., and Culbertson, M.R. (1991). The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon. Genes Dev. *5*, 2303–2314.

Lejeune, F., Li, X., and Maquat, L.E. (2003). Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. Mol. Cell *12*, 675–687.

Lewis, B.P., Green, R.E., and Brenner, S.E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc. Natl. Acad. Sci. USA *100*, 189–192.

Losson, R., and Lacroute, F. (1979). Interference of nonsense mutations with eukaryotic messenger RNA stability. Proc. Natl. Acad. Sci. USA *76*, 5134–5137.

Mendell, J.T., Sharifi, N.A., Meyers, J.L., Martinez-Murillo, F., and Dietz, H.C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts

and mutes genomic noise. Nat. Genet. *36*, 1073–1078.

Metzstein, M.M., and Krasnow, M.A. (2006). Functions of the nonsense-mediated mRNA decay pathway in Drosophila development. PLoS Genet. *2*, e180.

Muhlrad, D., and Parker, R. (1999). Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance. Rna *5*, 1299–1307.

Nguyen, L.S., Jolly, L., Shoubridge, C., Chan, W.K., Huang, L., Laumonnier, F., Raynaud, M., Hackett, A., Field, M., Rodriguez, J., et al. (2012). Transcriptome profiling of UPF3B/NMD-deficient lymphoblastoid cells from patients with various forms of intellectual disability. Mol. Psychiatry *17*, 1103–1115.

Okada-Katsuhata, Y., Yamashita, A., Kutsuzawa, K., Izumi, N., Hirahara, F., and Ohno, S. (2011). N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. Nucl. Acids Res. 1–16.

Peltz, S.W., Brown, A.H., and Jacobson, A. (1993). mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor. Genes Dev. *7*, 1737–1754.

Ramani, A.K., Nelson, A.C., Kapranov, P., Bell, I., Gingeras, T.R., and Fraser, A.G. (2009). High resolution transcriptome maps for wild-type and nonsense-mediated decay-defective Caenorhabditis elegans. Genome Biol. *10*, R101.

Rehwinkel, J., Letunic, I., Raes, J., Bork, P., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. Rna *11*, 1530–1544.

Singh, G., Rebbapragada, I., and Lykke-Andersen, J. (2008). A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. PLoS Biol. *6*, e111.

Takeshita, D., Zenno, S., Lee, W.C., Saigo, K., and Tanokura, M. (2007). Crystal structure of the PIN domain of human telomerase-associated protein EST1A. Proteins *68*, 980–989.

Weischenfeldt, J., Damgaard, I., Bryder, D., Theilgaard-Mönch, K., Thoren, L.A., Nielsen, F.C., Jacobsen, S.E.W., Nerlov, C., and Porse, B.T. (2008). NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. Genes Dev. *22*, 1381–1396.

Wittkopp, N., Huntzinger, E., Weiler, C., Sauliere, J., Schmidt, S., Sonawane, M., and Izaurralde, E. (2009). Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic development and survival. Mol. Cell. Biol. *29*, 3517–3528.

Wittmann, J., Hol, E.M., and Jäck, H.-M. (2006). hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay. Mol. Cell. Biol. *26*, 1272–

1287.

Xu, J., Kiel, M.C., Golshani, A., Chosay, J.G., Aoki, H., and Ganoza, M.C. (2006). Molecular localization of a ribosome-dependent ATPase on Escherichia coli ribosomes. Nucl. Acids Res. *34*, 1158–1165.

Yamashita, A., Ohnishi, T., Kashima, I., Taya, Y., and Ohno, S. (2001). Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay. Genes Dev. *15*, 2215–2228.

Zhang, S., Ruiz-Echevarría, M.J., Quan, Y., and Peltz, S.W. (1995). Identification and characterization of a sequence motif involved in nonsense-mediated mRNA decay. Mol. Cell. Biol. *15*, 2231–2244.